

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
BIOTEHNOLOOGIA ÕPPETOOL

**Meetodite võrdlus gRNA efektiivsuse hindamiseks genoomi muundamisel CRISPR/*Cas9*
tehnoloogiat kasutades**

Magistritöö

30 EAP

Rita Kreevan

Juhendaja Tõnis Org PhD

TARTU 2020

Meetodite võrdlus gRNA efektiivsuse hindamiseks genoomi muundamisel CRISPR/*Cas9* tehnoloogiat kasutades

CRISPR/*Cas9* eksperimendi efektiivsust mõjutab gRNA efektiivsus, mis juhib *Cas9* valgu sihtmärkjärjestuseni, ning seetõttu on oluline hinnata disainitud gRNA-de efektiivsust genoomi muundamisel. Kirjanduse põhjal populaarseima meetodi T7 endonukleas I tulemuste põhjal on raske hinnata gRNA tegelikku efektiivsust rakupopulatsioonis. Töös testitud kolm PCR-põhist meetodit: ACT-PCR, Bindel-PCR ja ORNi PCR ei andnud samuti rahuldavaid tulemusi. Nende meetodite puhul alahinnatakse populatsioonis olevate indelite osakaalu, sest rakupopulatsioonid sisaldavad ka muundamata järjestusi. Kuigi Sanger sekveneerimisel põhinevate meetoditega saab üldiselt edukalt hinnata indelite osakaalu ja gRNA efektiivsust, sõltuvad tulemused konkreetsest katsest. Kõige täpsema ja ülevaatlikuma pildi gRNA-de efektiivsusest ja tekkinud indelite sagedusest rakupopulatsioonis annab järgmise põlvkonna sekveneerimine.

Märksõnad: CRISPR/*Cas9*, gRNA

CERCS kood: B220, Geneetika, tsütogeneetika, T490, Biotehnoloogia

Comparison of methods to evaluate gRNA efficiency using CRISPR/*Cas9* technology for genome editing

The efficacy of CRISPR/*Cas9* is significantly affected by the designed gRNA that directs the *Cas9* protein to the target sequence. An important step in the CRISPR/*Cas9* experiment is to confirm the efficiency of designed gRNAs. Based on the literature, the most popular method - T7 endonuclease I gives vague results that do not accurately reflect gRNA efficiency. The three PCR-based methods tested in this work: ACT-PCR, Bindel-PCR and ORN PCR also did not give satisfactory results. These methods generally underestimate the proportion of indels because cell populations also often contain a higher proportion of wild-type sequences. Although, methods based on Sanger sequencing successfully assess the proportion of indels and gRNA efficiency, the results depend on the quality of the data. Next-generation sequencing provides the most accurate and comprehensive information about the efficiency of gRNAs and the frequency of indels in the whole edited cell populations.

Key words: CRISPR/*Cas9*, gRNA

CERCS code: B220, Genetics, cytogenetics, T490, Biotechnology

SISUKORD

KASUTATUD LÜHENDID	5
Sissejuhatus	6
1. Kirjanduse ülevaade	7
1.1 Genoomi muundamine.....	7
1.2 CRISPR/ <i>Cas</i> süsteem ja selle bioloogiline mehhanism.....	9
1.2.1 CRISPR/ <i>Cas9</i> kasutamine genoomi muundamises.....	11
1.3 CRISPR/ <i>Cas9</i> süsteemi ettevalmistamine ja rakku viimine.....	12
1.4 CRISPR/ <i>Cas9</i> gRNA efektiivsuse valideerimise meetodid	13
1.4.1 PCR-i põhised meetodid	14
1.4.1.1 Alleel-spetsiifiliste PCR-ide variatsioonid	14
1.4.1.1.1 ACT – PCR (<i>Annealing at critical temperature</i> PCR)	15
1.4.1.1.2 Bindel-PCR (<i>Biallelic indel</i> PCR)	16
1.4.1.1.3 ORNi-PCR (<i>Oligoribonucleotide interference</i> PCR)	17
1.4.2 Heteroduplekside analüüsimisel põhinevad meetodid	18
1.4.3 Sanger sekveneerimise ja järgmise põlvkonna sekveneerimine	19
2. EKSPERIMENTAALNE OSA	21
2.1 Töö eesmärgid.....	21
2.2 Materjal ja meetoodika	21
2.2.1 Rakuliin	21
2.2.2 Bakteritüvi.....	22
2.2.3 gRNAd sisaldavate CRISPR/ <i>Cas9</i> plasmiidide konstrueerimine	22
2.2.4 NIH3T6 rakkude transfektsioon CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 plasmiididega	24
2.2.5 Transfektsiooni efektiivsuse kontroll ja rakkude sorteerimine	24
2.2.6 DNA eraldamine	25
2.2.7 Illumina järgmise põlvkonna sekveneerimine.....	25
2.2.8 Sanger sekveneerimine.....	25
2.2.9 ACT-PCR.....	25
2.2.10 Bindel-PCR	26

2.2.11 ORNi-PCR	27
2.2.12 Illumina NGS ja Sanger sekveneerimise andmete analüüs	27
2.3 Tulemused.....	28
2.3.1 Plasmiidide valmistamine	28
2.3.2 NIH3T6 rakuliini transfekterimine CRISPR-Lyl1, CRISPR-Lyl2 ja CRISPR-Tal1 plasmiidiga	29
2.3.3 CRISPR-Tal1, CRISPR-Lyl1, CRISPR-Lyl2 segapopulatsioonide ja klonaalsete rakkude NGS tulemused.....	31
2.3.4 CRISPR-Tal1, CRISPR-Lyl1, CRISPR-Lyl2 segapopulatsioonide ja klonaalsete rakkude Sanger sekveneerimise tulemused.....	37
2.3.5 ACT-PCRi tulemused	41
2.3.6 Bindel-PCRi tulemused.....	43
2.3.7 ORNi PCR.....	44
2.4 Arutelu	46
KOKKUVÕTE	51
KIRJANDUSE LOETELU.....	53
LISAD	59
Lihtlitsents	67

KASUTATUD LÜHENDID

CRISPR - klasterdatud regulaarsete vahedega lühikesed palindroomsed kordused (ingl k *Clustered Regularly Interspaced Short Palindromic Repeats*)

Cas9 - CRISPR-ga seotud valk 9 (ingl k *CRISPR associated protein 9*)

crRNA – CRISPR RNA

DSB - kaheaheelaline katke (ingl k *Double Stranded Break*)

HDR - homoloogia poolt suunatud parandamine (ingl k *Homology-Directed Repair*)

Indel - insertioonid ja/või deletsioonid (ingl k *INsertions/DELetions*)

gRNA - giid-RNA (ingl *Guide RNA*)

NGS - järgmise põlvkonna sekveneerimine (ingl k *Next Generation Sequencing*)

NHEJ - mittehomoloogne otsade ühendamise (ingl k *Non-Homologus End Joining*)

PAM - protospesiseri külgnev motiiv (ingl k *Protospacer Adjacent Motif*)

TALEN - transkriptsiooni aktivaatori taoline efektor-nukleaas (ingl k *Transcription Activator-like Effector Nuclease*)

tracrRNA - trans-aktiveeruv CRISPR RNA (ingl k *Trans-activating crispr RNA*)

wt - metsik-tüüpi (ingl k *Wilde Type*)

ZFN - tsink-sõrme nukleaas (ingl k *Zinc Finger Nuclease*)

SISSEJUHATUS

Edusammud genoomi muundamise tehnoloogiates võimaldavad üha kiiremini ja lihtsamini muundada soovitud positsioone genoomis. Tänapäeval üks populaarseim genoomi muundamise tehnoloogia, CRISPR/*Cas9*, on avardanud meie teadmisi haigusseoselistest muutustest, genoomi organisatsioonist ning geenide omavahelistest interaktsioonidest. Võrreldes varasemalt kasutuses olevate tehnoloogiatega on CRISPR/*Cas9* kiirem, odavam ja lihtsamini kohandatav. *Cas9* poolt tekitatud kaheaheelaline katke parandatakse kas vigaderohke mitte-homoloogse otsade ühendamise (ingl k *Non-Homologous End Joining* – NHEJ või homoloogia poolt suunatud parandamisega (ingl k *Homology-Directed Repair* - HDR). Soovitud geeni funktsiooni uurimiseks kasutatakse ära eelkõige NHEJ vigaderohket olemust, kuna parandamisel tekkinud indelid viivad sageli geeni funktsiooni rikkumiseni. Üldiselt testitakse mitut kandidaat gRNA-d, kuna nende efektiivsus *Cas9* sihtmärkjärjestuseni juhtimisel on erinev. Seepärast on oluline valida testitavate gRNA-de seast kõige efektiivsem. gRNA-de efektiivsuse hindamiseks on mitmeid erinevaid meetodeid, aga neil kõigil on omad tugevad ja nõrgad küljed, mistõttu arendatakse pidevalt uusi meetodeid. Valitud meetodite juures on oluline, et need oleksid võimalikult lihtsad ja kiired ning tulemusi oleks võimalikult üheselt õigesti hinnata.

Kuna analüüsi meetodite „kuldstandardit“, järgmise põlvkonna sekveneerimist peetakse ajakulukaks, kalliks ning raskesti kättesaadavaks, püütakse arendada kiireid, lihtsaid ja kulutõhusaid meetodeid, mis võimaldaksid hinnata gRNA-de efektiivsust. Järjest rohkem on kirjanduses näha PCR-põhiste meetodite kohandamist gRNA efektiivsuse hindamiseks. Eelkõige põhinevad need meetodid lihtsal alleel-spetsiifilise PCR-i põhimõttel, mis on kombineeritud kas praimerite sulamistemperatuuri, MgCl₂ kontsentratsiooniga, fluorestseeruvate sondide või metsik-tüüpi järjestusi blokeerivate praimeritega. Kuna kirjanduses puuduvad artiklid, kus oleks hinnatud nimetatud meetodite efektiivsust, sai selle töö eesmärgiks testida kolme PCR-põhist meetodit ning võrrelda neid traditsioonilise Sanger sekveneerimise ja kaasaegsema järgmise põlvkonna sekveneerimisega. Testitud meetodite tulemuste võrdlemisel peaks selguma, millist informatsiooni on võimalik antud meetoditest saada ning kas lihtsus, kiirus ja kulude kokkuhoid on antud meetodite puhul põhjendatud, et valida neid järgmise põlvkonna sekveneerimise asemel.

1. KIRJANDUSE ÜLEVAADE

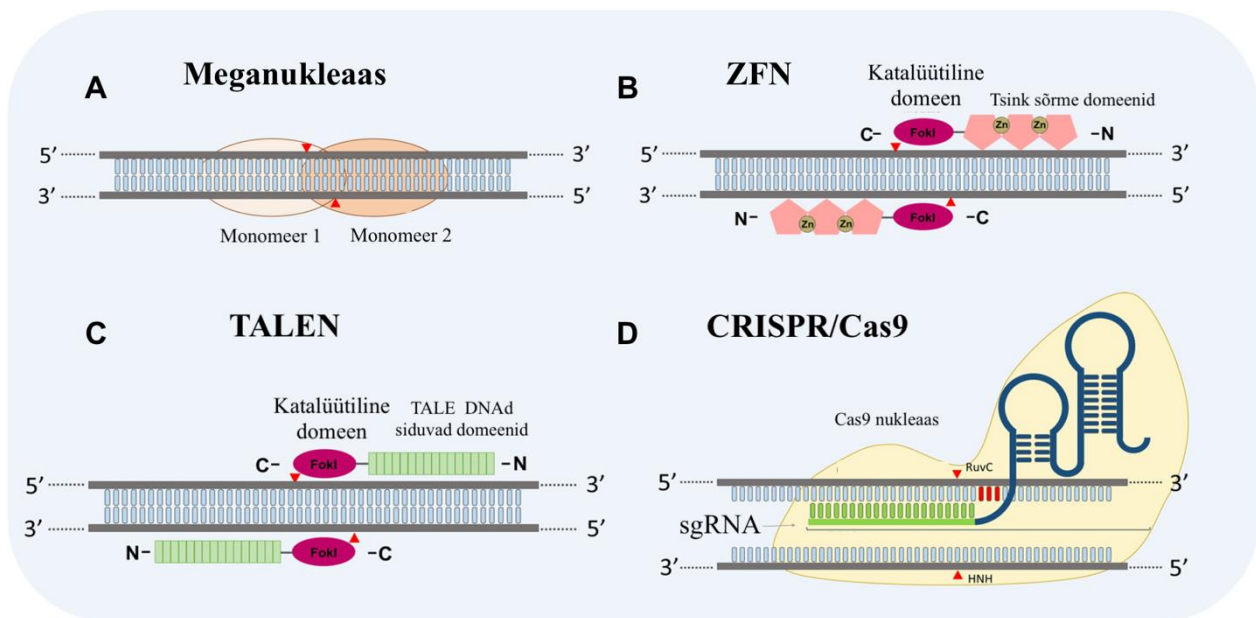
1.1 Genoomi muundamine

Viimastel aastatel on spetsiifiliste DNA järjestuste vastu suunatud modifitseeritud nukleas-ensüümid saanud märkimisväärset tähelepanu kui võimsad vahendid rakkude ja organismide geneetiliseks manipuleerimiseks, võimaldades sihtmärkgeeni deleteerimist, asendamist ja parandamist. Tänapäeval on mitmeid võimalusi (Tabel 1), kuidas erinevates organismides muuta soovitud positsiooni genoomis, kasutades selleks tsink-sõrme nukleaaase (ingl k *Zinc Finger Nuclease* – ZFN), meganukleaaase, transkriptsiooni aktivaatori taolisi efektor-nukleaaase (ingl k *Transcription Activator-like Effector Nucleases* – TALEN) ja klasterdatud regulaarsete vahedega lühikesi palindroomseid kordusi (ingl k *Clustered Regularly Interspaced Short Palindromic Repeats* – CRISPR) koos *Cas* valguga (Joonis 1).

Tabel 1. Programmeeritavad nukleasid genoomi muundamises

	Meganukleasid	ZFN	TALEN	<i>Cas9</i>
Sihtmärk järjestus (bp)	14-40	9-18	14-30	~23
Kui palju on sihtmärk järjestusi	Limiteeritud	Palju	Palju	Palju
Kohandamine	Keerukas	Keerukas	Lihtne	Väga lihtne
Suurus (kb)	~1	~1	~3	~3.5-4.5
Ära tundmine	Valk-DNA	Valk-DNA	Valk-DNA	RNA-DNA

Kõik ülalnimetatud meetodid töötavad sarnasel põhimõttel – nukleas suunatakse genoomis soovitud kohta, kus ta tekitab kaheaahelalise katke (ingl k *Double Stranded Break* – DSB). Edasi kasutatakse ära rakkudes olevaid DNA parandamise mehhanisme, mis käivituvad siis, kui DNA-s on tekkinud kaheaaheline katke. Kaheaahelised katked parandatakse kas mitte-homoloogse otsade ühendamise või, doonorjärjestuse olemasolul, homoloogia poolt suunatud parandamisega (Joonis 2) (Chandrasegaran ja Carroll, 2016; Gupta ja Musunuru, 2014). Nendest esimene, NHEJ, on vigaderohke ja tekitab parandamisel insertioone ja/või deletsioone (ingl k *INsertions/DELetions* - indel), mis võivad rikkuda geeni funktsiooni. Selliselt on võimalik saavutada huvipakkuva geeni *knock-out* ja uurida geeni funktsiooni erinevates rakuliinides või organismides.

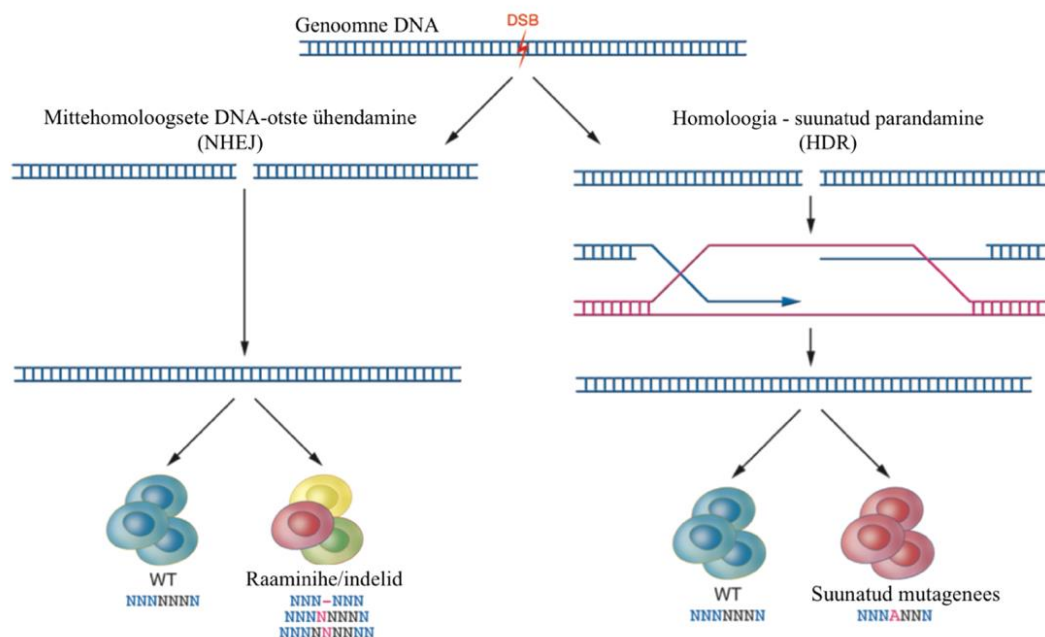


Joonis 1. Erinevate genoomi muundamiseks kasutatavate nukleaaaside ehitus. A) Meganukleaaasi homodimeerne struktuur. **B)** ZFN monomeerid, mis tunnevad ära sihtmärkjärjestuse on seotud *FokI* DNA-d lõikava domeeniga. **C)** Sarnaselt ZFN-ile sisaldab ka TALEN *FokI* DNA-d lõikavat domeeni, millele on liidetud TALE DNA-d siduvad domeenid. **D)** CRISPR/*Cas9* süsteem koosneb tüüpiliselt *Cas9* valgust, millel on kaks nukleaaasi domeeni RuvC ja HNH, ning kimäärsest giid-RNA-st (ingl k *guide RNA* – sgRNA) (Romay ja Bragard, 2017).

Meganukleasid on looduslikult esinevad restriksiooni ensüümid, mis on võimelised ära tundma suhteliselt pikki DNA järjestusi (14 – 40 bp) (Choulika *et al.*, 1995). Nende kohandamine soovitud sihtmärgi jaoks on keeruline, kuna DNA-d siduv domeen ja lõikav osa on põimunud üheks domeeniks. Seetõttu ei ole antud meetod leidnud laialdaselt kasutust (Silva *et al.*, 2011).

Uuema põlvkonna tehnoloogiad nagu ZNF-id ja TALEN-id on kimäärseid nukleasid, milles on ühendatud DNA-d äratundev domeen ja DNA-d lõikav *FokI*-nukleas domeen (Hensel ja Kumlehn, 2019; Wolfe *et al.*, 2000; Wood *et al.*, 2011). Kuna ZNF-des ja TALE-des on DNA-d siduv domeen eraldiseisev üksus *FokI* DNA-d lõikavast domeenist, on nende nukleaaaside kohandamine soovitud sihtmärgi jaoks oluliselt lihtsam.

Erinevalt valk-DNA interaktsioonidel põhinevatest meetoditest toimub CRISPR/*Cas9* kompleksi DNA-ga seondumine RNA vahendusel. Lühike gRNA molekul, mis on komplementaarne sihtmärkjärjestusega, juhib *Cas*-nukleaaasi sihtmärgini, misjärel *Cas*-valk teeb antud positsiooni kaheaheelise lõike. RNA vahendusel töötav süsteem lahendab mitmed probleemid: lihtsus (gRNA-d on lihtne disainida), kiirus ja hind.

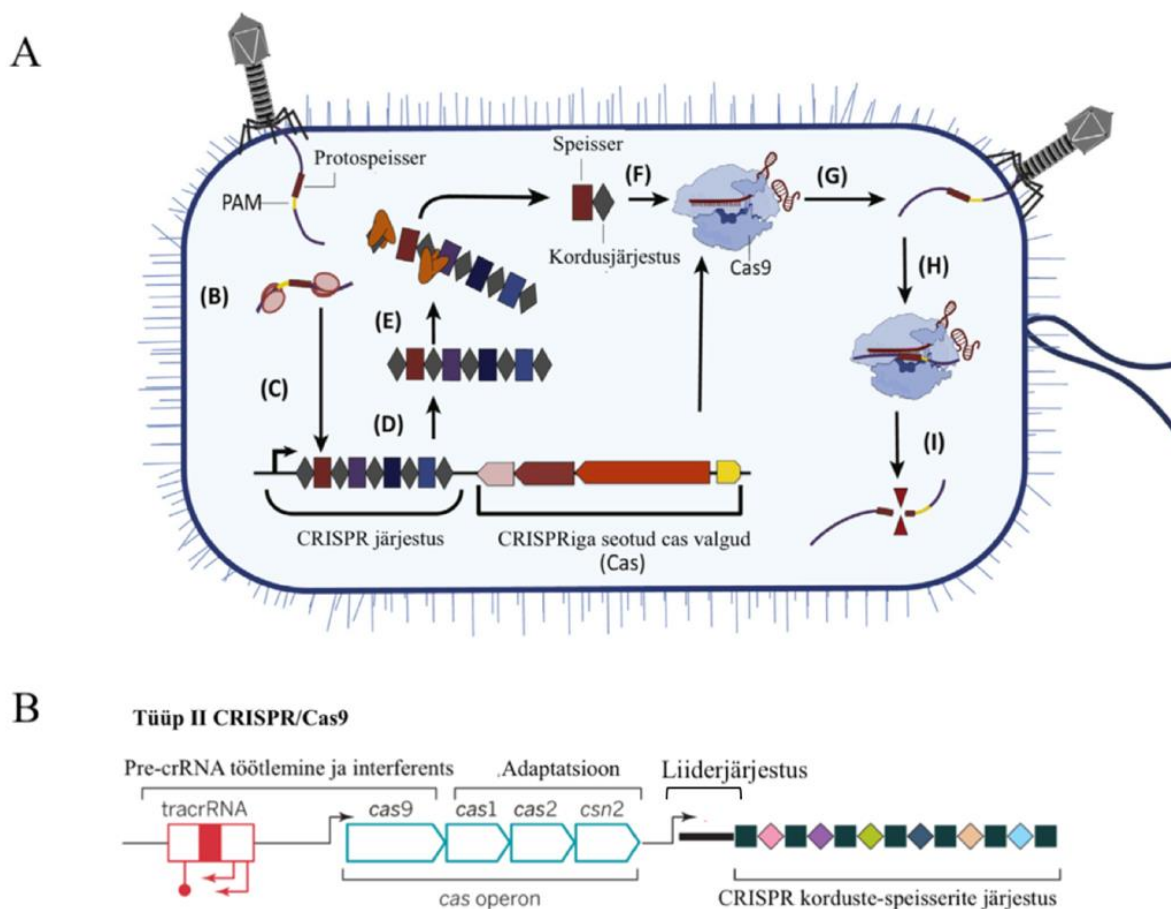


Joonis 2. Kaheaheelaliste katkete parandamise võimalused. Kaheaheelalise katke tagajärjel käivitub rakus üks kahest reparatsioonimehhanismist, juhul kui ei ole doonorjärjestust, mille järgi saab kaheaheelalist katket parandada, kasutatakse rakus NHEJ reparatsiooni, mis on oma loomult vigaderohke ja tekitab parandamisel erineva suurusega indeleid (joonisel vasakul), doonorjärjestuse olemasolul on võimalik kasutada homoloogia poolt suunatud parandamist (HDR) (joonisel paremal). Selliselt saab viia soovitud positsiooni huvipakkuva järjestuse, parandades näiteks vigase geeni (kohandatud Gupta ja Musunuru, 2014).

1.2 CRISPR/Cas süsteem ja selle bioloogiline mehhanism

CRISPR ehk klasterdatud regulaarsete vahedega lühikesed palindroomsed kordused on kindlasti olnud selle kümnendi üks olulisemaid arendusi biotehnoloogias. Tegemist on mikroobides avastatud adaptiivse immuunsüsteemiga, mis tänaseks on kohandatud lihtsaks ja kiireks genoomi muundamise tehnoloogiaks, võimaldades teha soovitud muudatusi genoomi erinevates positsioonides (Gupta *et al.*, 2019; Khanzadi ja Khan, 2019). CRISPR/Cas-iga saab sihtmärkjärjestust muuta oluliselt täpsemalt ja odavamalt kui ühegi teise senise lahendusega. Looduses on CRISPR-i kordusjärjestustel oluline osa mikroobide immuunsüsteemis, andes mikroobidele kaitse bakteriofaagide vastu (Joonis 3A) (Deveau *et al.*, 2010; Gupta *et al.*, 2019; Makarova *et al.*, 2011). Kui bakter puutub kokku bakteriofaagidega, rakendub mikroobides CRISPR/Cas süsteem, milles CRISPR-iga assotsieerunud valgud (ingl k *CRISPR associated protein* - cas) lõikavad võõra pärilikkusaine juppideks ning need lisatakse CRISPR kordusjärjestuste vahele (Joonis 3B) (Arslan *et al.*, 2014). Võõra pärilikkusaine fragmentidest

luuakse raamatukogu CRISPR kordusjärjestuste vahel, mis annab kaitse sama viiruse vastu ka tulevikus (Karvelis *et al.*, 2013; Wiedenheft *et al.*, 2012). Uuesti nakatudes transkribeeritakse speisser-kordusjärjestused *crRNA*-ks, mis koos *tracrRNA* ja *Cas*-valguga moodustavad efektorkompleksi. *Cas*-valk juhatakse *tracrRNA:crRNA* kompleksi poolt sihtmärgini, misjärel *Cas*-valk lõikab sihtmärkjärjestuse katki.



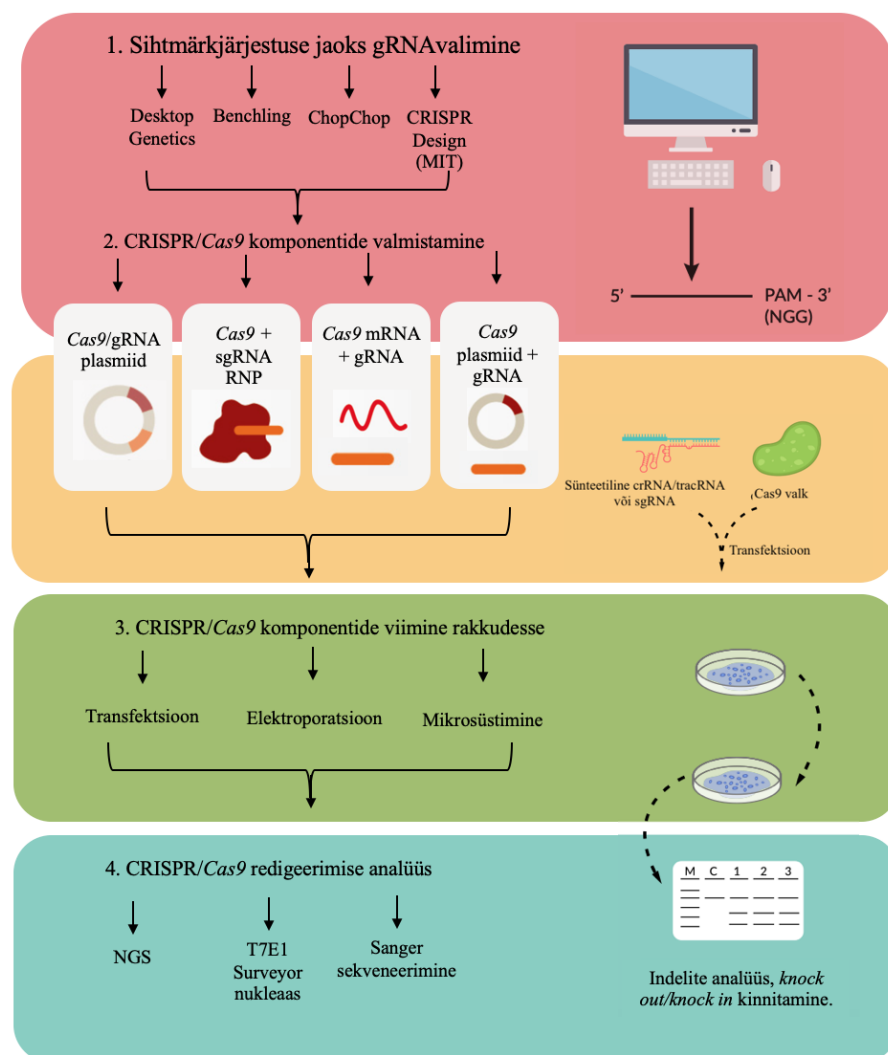
Joonis 3. CRISPR/Cas vahendatud immuunsus ja CRISPR/Cas lookuse ehitus. A) CRISPR/*Cas* vahendatud immuunsus algab infektsiooniga, kus bakterirakku siseneb bakteriofaagi pärilikkusaine. Bakteris rakenduvad *cas* valgud (**B**), mis seonduvad võõra pärilikkusainega (protospeisserid) (**C**), lisades need järjestused CRISPR kordusjärjestuste vahele. Uue nakatumise korral toimub (**D**) CRISPR järjestuste ekspressioon, mille käigus CRISPR järjestused töödeldakse *pre-crRNA*-deks. (**E**) Endonukleaaside abil töödeldakse speisser-korduselemendid küpseteks *crRNA*-deks. (**F**) *crRNA* seondues *Cas* valguga moodustab *cas:crRNA* kompleksi, mis otsib (**G**) komplementaarse sihtmärkjärjestuse, (**H**) kompleksi seondumisel sihtmärkjärjestusega (**I**) teeb *Cas*-valk võõr-DNA-sse kaheaahelalise katke, hoides ära edasise infektsiooni (Donohoue *et al.*, 2018). **B)** Tüüpiline CRISPR/*Cas* lookus koosneb *Cas* geenidest, mis aitavad läbi viia uute speisserite lisamist, *pre-crRNA* töötlemist ja võõr-DNA vaigistamist. Konserveerunud *Cas* geenideks on *Cas1* ja *Cas2*, mis esinevad tüüpiliselt

kõikides CRISPR süsteemides. Vahetult enne esimest CRISPR kordust asub AT-rikas liiderjärjestus, mille järjestus ja pikkus võib liigiti erineda. Palindroomsete CRISPR korduste vahele on lisatud eksogeensest DNA-st pärit speisserid, mille järjestuse alusel juhitakse *Cas*-valk komplementaarse sihtmärkjärjestuseni. Bakteriofaagide pärilikkusaine juppidest luuakse raamatukogu, mis sisaldab infot varasemalt kokku puutunud viirustega (Doudna ja Charpentier, 2014; Sorek *et al.*, 2008).

Pealtnäha universaalses süsteemis esineb siiski CRISPR-i tüüpide vahel erinevusi, läbi mille saavutatakse sihtmärgi ära tundmine, *pre-crRNA*-de töötlemine ja võõra pärilikkusaine vaigistamine. Hetkel kasutatakse genoomi muundamiseks enim tüüp II CRISPR süsteemi, eriti *Streptococcus pyogenes*-sest leitud *Cas9* valku (Shabbir *et al.*, 2016). Kuna antud süsteem vajab gRNA vahendatud võõr-DNA vaigistamiseks ainult ühte valku - *Cas9*, on see omadus osutunud väga kasulikuks just genoomi muundamise seisukohalt. Lihtne kasutamine ja paindlik kohandamine võimaldab CRISPR-it kasutada erinevates valdkondades ning erinevate organismide puhul.

1.2.1 CRISPR/*Cas9* kasutamine genoomi muundamises

Populaarseima CRISPR/*Cas9* meetodi võlu peitub tema lihtsuses, mis on saavutatud tänu kimäärsele gRNA-le, mida on lihtne ja odav disainida iga sihtmärgi jaoks. CRISPR/*Cas9* genoomi muundamise süsteem on tehnilikult valmistatud versioon *Cas9* valgust ja kimäärsest gRNA-st, mis koosneb *crRNA* ja *tracrRNA* liit-transkriptist (Jinek *et al.*, 2014). Soovitud kohas mutatsiooni tekitamiseks on vaja disainida sobilik gRNA (Joonis 4), mis kompleksis *Cas9*-ga juhib valgu sihtmärkjärjestuseni, misjärel *Cas9* saab teha järjestusse kaheaahelalise lõike. Lõike tegemiseks peab valitud kohale eelnema 3 nukleotiidiline PAM motiiv (ingl k *Protospacer Adjacent Motif* - PAM), mis *Cas*-valkude puhul erineb, kuid *Cas9* puhul on see 5'-NGG-3' (kus „N“ võib olla üks neljast nukleotiidist) (Hu *et al.*, 2018). CRISPR/*Cas9* tekitatud kaheaahelaline lõige parandatakse üldiselt NHEJ-ga või doonor-järjestuse olemasolul HDR-iga. NHEJ on enamasti aktiivne terve rakutsükli vältel ja on eelistatud DSB-te parandamise mehhanism, kuid sellega kaasnevad tihti erineva suurusega indelid, mis võivad viia geeni funktsiooni rikkumiseni (Kosicki *et al.*, 2018).



Joonis 4. Skemaatiline töövoog CRISPR/Cas9 ettevalmistamisest tulemuste analüüsini.

1.3 CRISPR/Cas9 süsteemi ettevalmistamine ja rakku viimine

Laboris kasutatav CRISPR/Cas9 koosneb sihtmärkjärjestuse spetsiifilisest gRNA-st ja Cas9 nukleasist. Genoomi efektiivseks muundamiseks on vaja, et mõlemad komponendid oleks sihtmärk-rakkudes ekspresseeritud (Robb, 2019). CRISPR/Cas9 kasutamine genoomi muundamiseks algab sobiliku gRNA disainimisega soovitud sihtmärgi jaoks (Joonis 4). gRNA disainimiseks on mitmeid programme, mis aitavad valida sobiliku gRNA järjestuse, millel oleks võimalikult vähe sihtmärgiväliseid efekte genoomis – *Benchling* (Benchling [CRISPR Guide RNA Design].2020), *ChopChop* (Labun *et al.*, 2019), *Desktop Genetics* (Hough *et al.*, 2016) jpt. Kuna Cas9 lõikekoha määrab ära 20-nukleotiidiline järjestus gRNA-s on *S. pyogenes*-es töötava CRISPR/Cas9 süsteemi jaoks oluline, et sihtmärkjärjestus eelneks 5'-NGG PAM motiivile. KO

(ingl k *Knock-out* – KO) eksperimendi puhul peaks eeldatav lõike koht jääma sihtmärk-geeni kodeeriva ala algusesse, mis tagaks geeni funktsiooni rikkumise (Ran *et al.*, 2013). Kuna gRNA-d töötavad erineva efektiivsusega, on mõistlik disainida mitu gRNA-d ühe sihtmärgi jaoks.

CRISPR/*Cas9* komponentide viimiseks rakkudesse kasutatakse kõige sagedamini plasmiidipõhist süsteemi (Wiles *et al.*, 2015). Ekspressiooniplasmiidi valmistamine on üldiselt lihtne ja kiire, sisaldades ühte kloonimisetappi oligonukleotiidiga. Sihtmärkjärjestuse jaoks disainitud oligonukleotiidid kloneeritakse plasmiidi, kus gRNA ekspressioon on näiteks U6 promootori all. 20 aluspaari pikkused oligonukleotiidid ligeeritakse plasmiidi, mis sisaldab *Cas9* kodeerivat järjestust ja gRNA jaoks vajaliku raamistikku (*scaffold*), mis vahendab *Cas9* ja gRNA seondumist (Ran *et al.*, 2013). Plasmide saab rakkudesse viia sobivate transfektsioonireagentidega, elektroporatsiooni või mikrosüstimisega. Raskemini transfekteeeritavate rakkude puhul saab plasmidi sisse viimiseks kasutada ka viirusvektoreid (Kabadi *et al.*, 2014). Üldjuhul sisaldab viirusvektori kasutamine keerulisemaid etappe, kuid on teatud rakutüüpide puhul efektiivsem kui plasmiidipõhine transfekteeerimine. Lisaks plasmidile saab *Cas9* nukleaasi ja gRNA-d rakkudesse viia RNA molekulidena (Liu *et al.*, 2017; Wilbie *et al.*, 2019). *Cas9* jaoks valmistatakse laboris *Cas9* valku kodeeriv mRNA ja gRNA valmistatakse *in vitro* transkriptsiooniga. Ka siin viiakse *Cas9* mRNA ja gRNA rakkudesse kas elektroporatsiooni või transfektsioonireagentidega. Alternatiivselt on võimalik kasutada *Cas9* ja gRNA ribonukleoalgu (ingl k *Ribonucleoprotein* – RNP) kompleksi (Gasiunas *et al.*, 2012), milles *Cas9* funktsionaalne valk ja gRNA on ühendatud üheks kompleksiks. Mõlemad komponendid on võimalik valmistada ise laboris või soovi korral osta valmiskujul. Oluline samm peale CRISPR/*Cas9* komponentide viimist rakkudesse on kontrollida CRISPR/*Cas9* süsteemi efektiivsust rakkudes, kas sihtmärkjärjestuses on tekkinud indeleid. Tänapäevaks on välja töötatud mitmeid meetodeid, mis võimaldaksid kiirelt ja lihtsalt kinnitada sihtmärkjärjestuses toimunud muutusi. Erinevatest meetoditest tuleb lähemalt juttu järgmises peatükis.

1.4 CRISPR/*Cas9* gRNA efektiivsuse valideerimise meetodid

Ligikaudu 40-60% rakkudest, kuhu on edukalt sisse viidud CRISPR/*Cas9* konstrukti, sisaldavad soovitud sihtmärkjärjestuses mutatsioone (Robb, 2019). Seega on vajadus tundlike ja usaldusväärsete meetodite järele, millega oleks võimalik hinnata CRISPR/*Cas9* lõikamise efektiivsust rakupopulatsioonis (Yu *et al.*, 2014). CRISPR-i genoomi muundamise efektiivsuse

analüüsiks olemasolevate meetodite valik sõltub sellest, millised vahendid on kellelegi kättesaadavad. Mutatsioonide tuvastamiseks on välja töötatud mitmeid erinevaid meetodeid (Lisa 1), sealhulgas erinevad PCR-põhised meetodid, restriksiooni ensüümide (ingl k *Restriction Enzyme assay* – RE) testid, heteroduplekside analüüsil põhinevad meetodid nagu T7 endonukleaasi I (T7EI) (Vouillot *et al.*, 2015) ja *Surveyor*-nukleaasi analüüs (Qiu *et al.*, 2004), polüakrüülamiid geelelektroforeesil (ingl k *Polyacrylamide Gel Electrophoresis* – PAGE) põhinevad meetodid, kõrge eraldusvõimega sulamise (ingl k *High-Resolution Melting Analyse* – HRMA) (Montgomery *et al.*, 2007; Samarut *et al.*, 2016) analüüsil põhinevad testid ja fluorents-PCR-kapillaargeeli elektroforeesi meetodid (Hua *et al.*, 2017; Montgomery *et al.*, 2007; Ramlee *et al.*, 2015, 2015; Zhu *et al.*, 2015). Paljud neist meetoditest on aeganõudvad, vajavad spetsiaalset aparatuuri ning oskusi ning seetõttu otsitakse pidevalt uusi meetodeid, mis oleksid töökindlad, lihtsad, odavad ja kiired.

1.4.1 PCR-i põhised meetodid

Lisas 1 välja toodud meetodid kasutavad esimese etapina PCR-i ja võivad seetõttu alahinnata sihtmärgipäraste (*on-target*) mutatsioonide sagedust, kuna deletsioonid võivad ulatuda ka väljapoole PCR-i amplikonidest (Kosicki *et al.*, 2018). Siiski on PCR-i põhised meetodid sobilikud esialgseks CRISPR/Cas9 gRNA efektiivsuse kinnitamiseks, kuna väikesed indelid tekivad suurema sagedusega kui suured (Chakrabarti *et al.*, 2019). Antud meetodite puhul tehakse esimese etapina PCR, mille käigus amplifitseeritakse sihtmärkjärjestus, mis peaks sisaldama erinevaid indeleid. Järgnevatel etappidel saab ära kasutada erinevaid vahendeid, millega testida mutatsioonide olemasolu. Näiteks saab kasutada mutatsiooni tekke koha spetsiifilisi praimereid, MgCl₂ kontsentratsiooni, renaturatsiooni temperatuuri ja mitmeid ensüüme (Hua *et al.*, 2017; Sakurai *et al.*, 2019; Vouillot *et al.*, 2015).

1.4.1.1 Alleel-spetsiifiliste PCR-ide variatsioonid

PCR on laialt kasutatav meetod, millega on võimalikult suhteliselt kiirelt ja lihtsalt testida suurel hulgal proove. Üks PCR-i tsükkel koosneb kolmest etapist: denaturatsioon, seondumine ja elongatsioon. PCR-i edukaks toimimiseks on olulised mitmed komponendid ja parameetrid nagu näiteks praimerite sulamistemperatuur, polümeraas ja selle kofaktor MgCl₂. Seondumistemperatuur on oluline praimeri sihtmärgile seondumiseks ja seeläbi ka PCR-i

õnnestumiseks. Optimaalne seondumistemperatuur hoiab ära ebaspetsiifilise seondumise, vähendades sellega mittespetsiifiliste produktide teket (Hua *et al.*, 2017). Magneesium on oluline polümeraasi seondumiseks ja DNA ahelate pikendamisel. Paljude mutatsioonide tuvastamiseks on ära kasutatud lihtsat alleel-spetsiifilist PCR-i (Ugozzoli ja Wallace, 1991). Mutatsiooni analüüsiks kasutatava alleel-spetsiifilise PCR-i puhul disainitakse praimerid, millega oleks võimalik eristada mutatsiooni sisaldavat alleeli metsik-tüüpi alleelist. Praimerid disainitakse selliselt, et eeldatav mutatsiooni positsioon jääks praimeri 3' otsa, kuna DNA-polümeraasi töö sõltub otseselt praimeri 3' otsa täpsest paardumisest sihtmärkjärjestusega. Praimerite seondumine sihtmärgile on stabiilne vaid siis, kui tegemist on täieliku komplementaarsusega, kuid indelite olemasolul ei seonu 3' ots sihtmärgiga ja amplifikatsiooni ei toimu (Ugozzoli ja Wallace, 1991). Selle meetodi põhimõttel on loodud mitmeid teste, mis võimaldaksid tuvastada CRISPR/Cas9 tekitatud indeleid kiirelt ja lihtsalt. Meetodid kombineerivad alleel-spetsiifilisi primereid kas renaturatsiooni temperatuuri muutmisega (Hua *et al.*, 2017), *nested*-PCR-iga, MgCl₂ kontsentratsiooniga (Sakurai *et al.*, 2019), fluorestseeruvate sondidega (Findlay *et al.*, 2016; Mock *et al.*, 2016; Ramlee *et al.*, 2015) või PCR-i blokeerivate oligonukleotiididega (Fujita *et al.*, 2018; Harayama ja Riezman, 2017).

1.4.1.1.1 ACT - PCR (*Annealing at critical temperature* PCR)

Praimerite seondumis-temperatuuri ja alleel-spetsiifilisi primereid on ära kasutatud sellise meetodi puhul nagu ACT-PCR (ingl k *Annealing at Critical Temperature* – ACT) (Hua *et al.*, 2017). Antud meetodi puhul leitakse kõrgeim temperatuur, mille juures toimub edukas PCR mutatsioonitundlike praimeritega. F praimer disainitakse kohakuti eeldatava Cas9 lõikamiskohaga, jättes kuni 4 bp üleulatuva otsa (Joonis 5). Indelite olemasolul ei ole F praimeri 3' ots võimeline kõrgemal temperatuuril sihtmärkjärjestusega seonduma ning amplifikatsiooni ei toimu.

A) Primerite disain

DSB
↓

F primer 4bp 5'-NNNNNNNNNNNNNNNNNNNN-3'

üleulatava otsaga

$T_m \text{ R primer} > T_m \text{ F primer}$

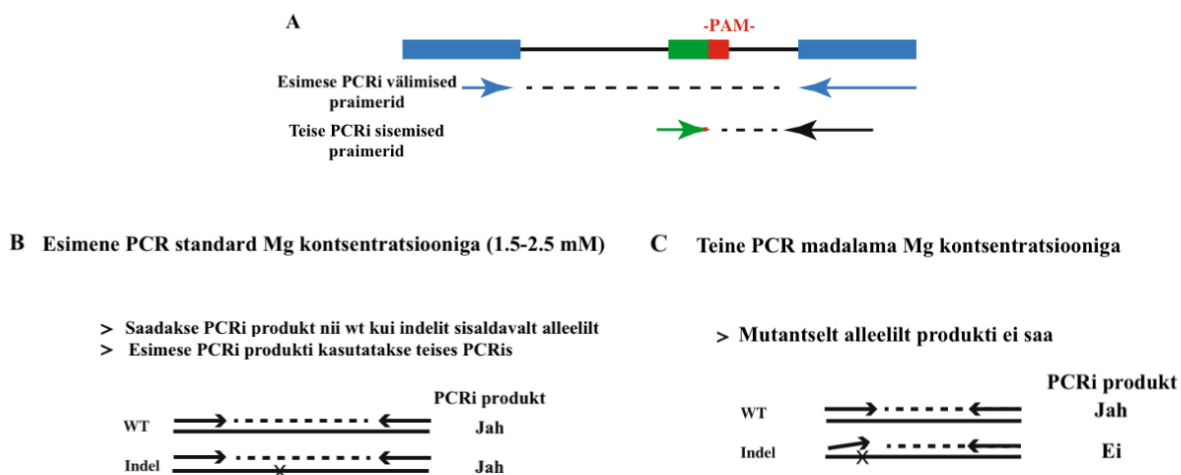
R primer 3'-NNNNNNNNNNNNNNNNNNNN-5'

The diagram illustrates the mechanism of a mismatch repair (MMR) system. It shows a double-stranded DNA molecule with a mismatch (indicated by a red 'X'). The process involves several steps: Denaturatsioo (strand separation), Seondumine (mismatch recognition), and Elongatsioo (mismatch repair). The final step is Amplifikatsioo (amplification). The diagram shows that the mismatch repair system successfully repairs the mismatch in the WT (Wild Type) but fails to do so in the Mutant, as indicated by the red 'X' remaining in the mutant strand.

1.4.1.1.2 Bindel-PCR (*Biallelic indel PCR*)

16

praimereid, millega saab sihtmärkjärjestust sisaldavad ampliconid. Teine PCR viiakse läbi madalamal $MgCl_2$ kontsentratsioonil, kasutades eeldatava mutatsiooni tekkekoha-spetsiifilisi sisemisi praimereid. F praimer on disainitud järjestusele, mille tunneb ära gRNA ning kuhu tehakse Cas9 poolt DSB ja R praimer on disainitud ~200 bp allavoolu PAM motiivist.



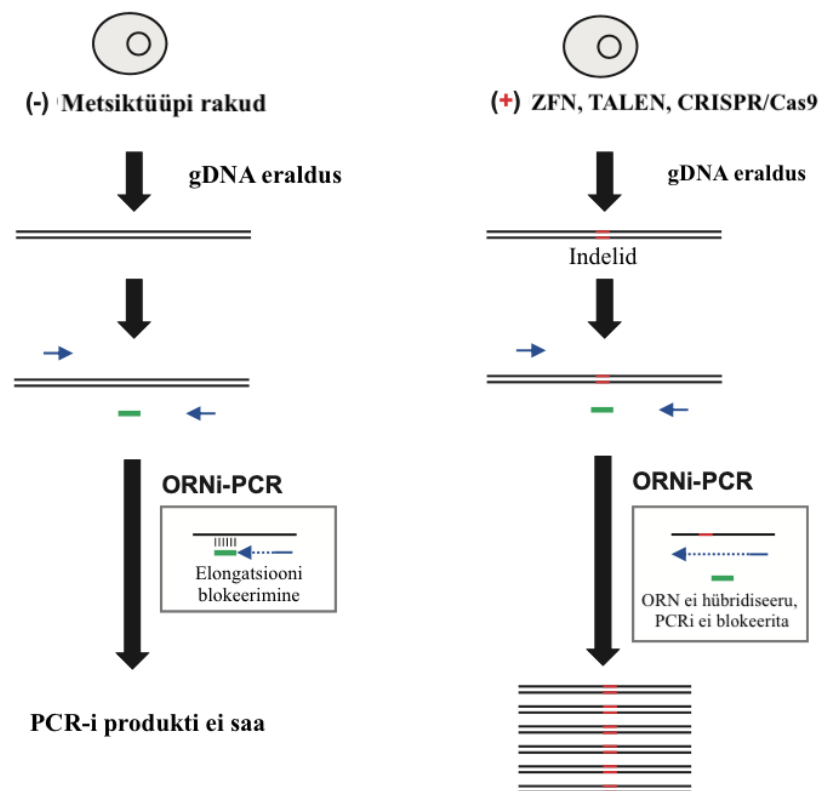
Joonis 6. Skemaatiline ülevaade Bindel-PCRist. A) PCRi jaoks disainitakse kaks paari praimereid: välimised praimerid esimese PCR-i jaoks ja sisemised mutatsioonitundlikud praimerid teise jaoks. B) Esimene PCR viiakse läbi standardtingimustel, kus kasutatakse 1.5-2.5 mM $MgCl_2$ kontsentratsiooni, saadakse produkt nii metsik-tüüpi kui indelit sisaldavalt alleelilt. C) Teine PCR viiakse läbi madalamal $MgCl_2$ kontsentratsioonil, kasutades esimesest PCR-ist saadud produkti. Madalama $MgCl_2$ kontsentratsioonil mutatsiooni sisaldavalt alleelilt enam PCR-i produkti ei saada (kohandatud Sakurai *et al.*, 2019).

Kasutades antud praimerite paari ja madalamat $MgCl_2$ kontsentratsiooni, ei seonu mutatsiooni tundlik F praimeri 3' ots indelite olemasolul sihtmärkjärjestusega ja amplifikatsiooni ei toimu. PCR-i produkti olemasolu viitab metsik-tüüpi DNA-le ja produkti puudumine kinnitab indelite esinemist järjestuses.

1.4.1.1.3 ORNi-PCR (*Oligoribonucleotide interference PCR*)

Oligoribonukleotiidi interferents (ingl k *OligoRiboNucleotide Interference* – ORNi) PCR-i puhul kasutatakse mutatsioonide kinnitamiseks PCR-i blokeerivaid oligoribonukleotiide (Fujita *et al.*, 2018). Selleks disainitakse sihtmärkjärjestusega komplementaarsed 17-29 bp ORN-id, mida kasutatakse PCR-i läbiviimisel. Metsik-tüüpi järjestuse puhul seondub ORN sihtmärkjärjestusega, blokeerides PCR-i, kuid mutatsiooni olemasolul ORN sihtmärkjärjestusega ei seonu ja toimub

sihtmärkjärjestuse amplifikatsioon (Joonis 7). Edukas PCR-i amplifikatsioon viitab sihtmärkjärjestuses olevale mutatsioonile ning PCR-i produkti puudumine viitab muundamata järjestusele.



Joonis 7. Skemaatiline ülevaade ORNi PCR-i tööpõhimõttest. Disainitakse *Cas9* eeldatava kaheahelalise lõike toimumiskohaga kattuvad ORN-id, mis metsik-tüüpi järjestuse puhul paarduvad sihtmärgiga ja blokeerivad PCR-i amplifikatsiooni. Mutatsiooni olemasolul ORN-id ei paardu sihtmärgiga ning toimub sihtmärkjärjestuse amplifikatsioon (kohandatud Fujita *et al.*, 2018).

1.4.2 Heteroduplekside analüüsimisel põhinevad meetodid

Populaarsed mutatsiooni tuvastamise meetodid on olnud ka erinevad heteroduplekside lõikamisel põhinevad testid nagu T7 endonukleas I (T7EI) test (Mashal *et al.*, 1993) ja *Surveyor*-nukleasi analüüs (Yang *et al.*, 2000). Need meetodid kasutavad ära asjaolu, et disainernukleasidega töödeldud rakkude populatsioonist eraldatud genoomne DNA on heterogeenne segu mutantsetest alleelidest ja metsik-tüüpi alleelidest. Testid põhinevad sihtmärk-lookuse amplifitseerimisel, kus renaturatsiooni etapis võivad kaks DNA fragmenti, metsik-tüüpi ja muteerunud fragment, omavahel seonduda. Valepaardumine kahe ahela vahel tekitab juuksenõela struktuuri, mida

lõigatakse nukleaaside poolt. Mida efektiivsem on CRISPR/*Cas9*, seda rohkem on indeleid ja heteroduplekside moodustumisi. Lõikamise järel saadud fragmendid erinevad suuruselt ja on visualiseeritavad geelelektroforeesiga (Vouillot *et al.*, 2015).

Võrreldes *Surveyor*-nukleaasiga on T7EI tundlikum just suuremate indelite tuvastamisel, aga ei suuda edukalt lõigata lühemaid indeleid ja üksiku nukleotiidi muutusi. T7EI nukleaasi efektiivsus sõltub inkubeerimise tingimustest, DNA/ensüümi vahekorra ja teistest reaktsiooni parameetritest, mis nõuavad optimeerimist. *Surveyor*-nukleaas on kallim ensüüm ja üldiselt madalama tundlikkusega, aga võimaldab tuvastada lühemaid indeleid ja üksikute nukleotiidide muutusi. Kuigi T7EI on üsna populaarne, annab see ainult ligikaudse hinnangu tekkinud mutatsioonide kohta. Üldiselt ei ole antud meetodite tulemused järjepidevad ja indelite sagedus ei ole kooskõlas sellega, mis on tegelik indelite tekkesagedus rakupopulatsioonis (Kim *et al.*, 2011).

Lisaks sobib heteroduplekside analüüsiks ka kõrge resolutsiooniga sulamiskurvi analüüs (ingl k *High Resolution Melting Analysis* – HRMA) (Samarut *et al.*, 2016). HRMA puhul mõõdetakse denaturatsiooni ajal fluorestsentsi kadumist, kui vabanevad kaheaheelalise DNA-ga interkaleeruvad värvid. Selliselt registreeritakse amplikonide temperatuurist sõltuv denaturatsiooni profiil, mis erineb mutatsiooni sisaldavate amplikonide ja metsik-tüüpi amplikonide vahel. Sulamistemperatuuri muutused ja sulamiskõverate kuju annavad mõlemad kasulikku teavet nii homosügootsete alleelvariantide esinemise kohta, kui ka heterosügootsete mutatsioonide kohta (Farrar ja Wittwer, 2017). Saavutamaks paremat tundlikkust, on selle meetodi puhul eelistatud lühemad amplikonid, kuna annavad suurema nihke sulamistemperatuuris. Üldiselt on HRMA kiire, analüüsi teostamine võtab aega <2 tundi ja kuna see meetod ei kahjusta saadud amplikone, on võimalik neid edasi analüüsida näiteks geelelektroforeesiga. Siiski on ka sellel meetodil omad miinused: saadud fragmendid on suhteliselt väiksed ja nii võivad jääda tuvastamata suuremad indelid. Lisaks on selle meetodi jaoks vaja spetsiifilist aparatuuri ja analüüsi tarkvara, mis mõlemad on kallid (Bell *et al.*, 2014; Montgomery *et al.*, 2007).

1.4.3 Sanger sekveneerimise ja järgmise põlvkonna sekveneerimine

CRISPR/*Cas9* tekitatud indelite määramiseks on võimalik kasutada ka sekveneerimisel põhinevaid meetodeid. Sanger sekveneerimine on olnud tänu lihtsale ja kuluefektiivsele töövoole ning lihtsale andmeanalüüsile eelistatud meetod indelite iseloomustamiseks (Brinkman *et al.*, 2014). Tänu arengutele analüüsi programmides, saab Sanger sekveneerimisega analüüsida ka rakkude

segapopulatsioonide (Brinkman ja van Steensel, 2019). Sekveneerimise meetodi eeliseks on saadud teave mutatsioonide tüübi ja sageduse kohta sihtmärkjärjestuses. Sanger sekveneerimisest saadud tulemuste analüüsimiseks on välja töötatud ka mitmeid veebipõhiseid andmeanalüüsi platvorme, mis pakuvad kiiret ja lihtsat andmeanalüüsi võimalust (Brinkman ja van Steensel, 2019; Bruyneel *et al.*, 2019; Park *et al.*, 2017; Pinello *et al.*, 2016).

Järgmise põlvkonna sekveneerimine on tundlik meetod ja annab seetõttu väga põhjaliku pildi tekkinud indelitest, nende suurusest ja sagedusest. NGS võimaldab detekteerida indeleid nii segapopulatsioonides kui kloonrakuliinides. Võimaldades analüüsida suurel hulgal proove, on NGS saanud populaarseks valikuks CRISPR/Cas9 tekitatud indelite hindamisel (Bell *et al.*, 2014; Ramlee *et al.*, 2015). Samas vähese proovide hulga korral on NGS kallis, mistõttu ei ole see mõistlik väiksemamahulistest uuringutes. Juhul kui laboril on ressursse, on see suurepärane meetod analüüsima CRISPR/Cas9 tekitatud muudatusi. NGS-i saadud toorandmete analüüsiks on praeguseks välja töötatud mitmeid erinevaid programme ja vastavalt oskustele on võimalik valida enda jaoks sobiv.

NGS-i ja Sanger sekveneerimise ühine eelis seisneb asjaolus, et nad on järjestuspõhised ja võimaldavad seega hinnangut lõikamise efektiivsuse kohta ja teavet indelite mustrite kohta. Selline teave on äärmiselt oluline, et analüüsida CRISPR/Cas9-ga saadud tulemusi.

2. EKSPERIMENTAALNE OSA

2.1 Töö eesmärgid

Antud magistritöö on osa projektist, mille üheks eesmärgiks oli Tal1 ja Lyl1 geenide *knock-out* rakuliinide loomine kasutades CRISPR/*Cas9* meetodit. Esialgsed katsed kirjanduse põhjal laialdasemalt kasutatud T7 endonukleas I meetodiga CRISPR/*Cas9* efektiivsuse hindamisel andsid mitterahuldavaid tulemusi. Sellest lähtuvalt sai antud magistritöö eksperimentaalse osa eesmärgiks välja selgitada sobivaim meetod, mis võimaldaks kinnitada CRISPR/*Cas9* jaoks disainitud gRNA-de efektiivsust genoomi muundamisel. Töö koosnes järgmistest etappidest:

1. Kirjanduse põhjal valiti välja 3 PCR-i põhist meetodit, mis lubavad kiirelt ja lihtsalt teha kindlaks CRISPR/*Cas9* muundatud rakkudes esinevaid indeleid. Lisaks analüüsiti proove ka Sangeri ja Illumina NGS sekveneerimise meetoditega.
2. Välja valitud meetodite testimiseks transfekteeeriti NIH3T6 rakke kahe geeni (Tal1 ja Lyl1) vastu suunatud kolme erineva gRNA (CRISPR-Lyl1, CRISPR-Lyl2 ja CRISPR-Tal1) CRISPR/*Cas9* konstruktiga.
3. Testiti valitud meetodite efektiivsust indelite kinnitamisel transfekteeeritud segarakupopulatsioonides ja sorteeritud klonaalsetes rakupopulatsioonides.

2.2 Materjal ja metoodika

2.2.1 Rakuliin

Antud töös kasutati hiire fibroblasti NIH3T6 rakuliini. Rakke kasvatati DMEM (*Dulbecco's Modified Eagle Medium*) (Naxo, Eesti) söötmes, millele oli lisatud 10% veise loote seerumit (ingl *Fetal Bovine Serum* - FBS) ning 1% penitsilliini ja streptomütsiini segu (Naxo, Eesti). Rakuliini kasvatati 37°C 5% CO₂ sisalduse juures MCO-19AIC inkubaatoris (Sanyo Electric Co, Jaapan). Rakkude tihedus määrati kindlaks hemotsütoomeetriga (Hausser Scientific, USA).

2.2.2 Bakteritüvi

Plasmiidide paljundamiseks kasutati kompetentseid *E. Coli* bakterirakke (Invitrogen™ One Shot™ TOP10 Chemically Competent *E. Coli*), mida kasvatati Lurian-Bertani (LB) vedelsöötmes (trüptoon 10 g/l, pärmiekstrakt 5 g/l, NaCl 10 g/l).

2.2.3 gRNAd sisaldavate CRISPR/Cas9 plasmiidide konstrueerimine

Soovitud plasmiidide loomiseks kasutati selgroona pX-458 *NotI* plasmiidi (Lisa 2). Sisestamaks antud plasmiidi *Lyl1* geeni vastu suunatud gRNA-d disainiti esmalt vastavad oligonukleotiidid, milleks kasutati ChopChop programmi (Labun *et al.*, 2019) (Tabel 2). gRNAd disainiti geeni algusesse, et tekkinud mutatsioonid viiksid suurema tõenäosusega mittefunktsionaalse geeni tekkimiseni. Geeni *Lyl1* vastu valmistati kaks erinevat gRNAd sisaldavat plasmiidi (CRISPR-*Lyl1* ja CRISPR-*Lyl2*). *Tal1* geeni vastu suunatud gRNAd sisaldava CRISPR-*Tal1* plasmiidi valmistas Mare Vahtre oma bakalaureusetöö raames (Vahtre, 2019).

Tabel 2. *Lyl1* ja *Tal1* oligonukleotiidid

<i>Tal1</i>	<i>Forward</i>	5' CACCCCTGTTACAGGATGACGGAG 3'
	<i>Reverse</i>	5' AAACCTCCGTCATCCTGTAACAGG 3'
<i>Lyl1</i>	<i>Forward</i>	5' CACCGGGCGGGCACATGGACCCCA 3'
	<i>Reverse</i>	5' AAACCTGGGGTCCATGTGCCCCGCCC 3'
<i>Lyl2</i>	<i>Forward</i>	5' CACCGGGGTCCATGTGCCCCGCCCC 3'
	<i>Reverse</i>	5' AAACGGGGCGGGCACATGGACCCCA 3'

pX-458 *NotI* plasmiid lõigati lahti *BbsI* restriksiooniensüümiga (Thermo Scientific, USA) tootjapoolse protokollil alusel: 2 µl plasmiidi, 2 µl *FastDigest BbsI* restriksiooniensüümi, 4 µl 10x *FastDigest* puhvrit, 30 µl mQ. Proove inkubeeriti 37°C 60 min. Lõigatud plasmiid puhastati geelelektroforeesi abil, 1% agarosi geelil 0,5x TAE puhvris (Tris/Borate/EDTA), millele lisati etiidumbromiidi (0,5µg/ml). Õige pikkusega DNA fragment lõigati agarosgeelilt välja ja puhastati *Macherey-Nagel NucleoSpin Gel and PCR Clean-up* komplektiga vastavalt tootja protokollile.

Disainitud oligonukleotiidid telliti firmast Metabion, renatureeriti tuubis kokku: 4 µl *forward* ja 4 µl *reverse* oligonukleotiidi (100 µM) segati kokku ja inkubeeriti PCR-i termotsükleris (*MJ Research PTC-200 Peltier Thermal Cycler*), kasutades järgnevat programmi: 37°C 30 min, 95°C 5 min, 95°C jahutades 0,1°C sekundis 25°C-ni. Oligonukleotiidide plasmidi ligeerimiseks valmistati reaktsioonisegu, mis sisaldas 1 µl 10x T4 ligaasi puhvrit, 0,25 µl T4 ligaasi, 1,5 µl plasmidi, 1 µl renatureeritud oligonukleotiide, 6,25 µl mQ. Segu hoiti toatemperatuuril 10 minutit. Transformatsiooniks kasutati 50 µl kompetentseid bakterirakke (*Invitrogen™ One Shot™ TOP10 Chemically Competent E. Coli*). Külmutatud bakterirakkude suspensiooni sulatati jääl 5 min. Bakteritele lisati 5 µl ligeeritud plasmiidset DNA-d ning segu inkubeeriti jääl 20 min, sellele järgnes kuumašokk 42°C 45 sekundit ning rakud asetati tagasi jääle. Pärast 1 ml LB söötme lisamist inkubeeriti rakke loksutil 37°C 1h juures ja tsentrifuugiti 5000x g juures 5 min. Supernatant eemaldati ja bakterirakud võeti ülesse 50 µl-is vedelsöötmes. Rakususpensioon külvati ampitsilliiniga LB-tardsöötmele ja jäeti kasvama 37°C juurde. 24h möödudes valiti selektiivse antibiootikumiga LB-tardsöötmele välja üksikud kolooniad, mida kasvatati sama antibiootikumiga 2 ml LB-vedelsöötmes loksutil üleöö.

Järgnevalt viidi läbi plasmiidide eraldamine bakterirakkudest *Macherey-Nagel NucleoBond Xtra Mini* komplektiga tootjapoolse protokollil alusel. Ettekasvatatud bakterirakud kanti üle 1,5 ml tuubi ning fuugiti 3000x g 10 minutit. Eemaldati supernatant ning rakud resuspendeeriti 300 µl-is resuspensiooni puhvrts, lisati 300 µl lüüsi lahust. Tuube kallutati järgnevalt 5 korda, misjärel lisati 300 µl neutraliseerimise lahust ja tuube kallutati kuni sinise värvi kadumiseni. Seejärel fuugiti tuube 13000x g juures 10 minutit. Eemaldati 750 µl supernatanti, mis viidi uude 1,5 ml tuubi ning sellele lisati 750 µl isopropanooli. Tuube fuugiti 13000x g 10 minutit, eemaldati supernatant ning eraldatud DNA-le lisati 500 µl 70% etanooli. Tuube fuugiti 13000x g juures 10 minutit, eemaldati supernatant ning tuube kuivatati 37°C juures avatud korgiga etanooli täieliku aurustumiseni. Seejärel lahustati eraldatud DNA 30 µl-is mQ-s. Eraldatud DNA kontsentratsioon mõõdeti *NanoDrop2000* spektrofotomeetriga (Thermo Scientific). Plasmide kontrolliti *NotI* restriksiooniensüümiga (Thermo Scientific) lõikamise abil, kasutades tootja poolt etteantud protokollil. Restriksioonisegu sisaldas 16 µl mQ, 2 µl 10x *Buffer* 1 µl DNA, 1 µl *NotI* restriksiooniensüümi. Reaktsioonisegu inkubeeriti 37°C juures 1h, millele järgnes geelelektroforees 1.5% agarosgeelil. Õigeid plasmide paljundati 200 ml LB vedelsöötmes, millele oli lisatud 200 µl 100 mg/ml kontsentratsiooniga ampitsilliini. Plasmidne DNA eraldati bakterirakkudest *Macherey – Nagel NucleoBond Xtra Midi* komplektiga tootjapoolse protokollil

alusel. Plasmiidide kontsentratsioon mõõdeti *Nanodrop2000* spektrofotomeetriga (Thermo Scientific) ja kontrolliti *NotI* restriksiooniensüümiga lõigates. Valmistatud plasmiidid kontrolliti üle ka Sanger sekveneerimisega Tartu Ülikooli Genoomika Instituudi sekveneerimise tuumiklaboris.

2.2.4 NIH3T6 rakkude transfektsioon CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 plasmiididega

NIH3T6 rakuliini transfektsiooniks kasutati jetPRIME® (JP) ja jetOPTIMUS® (JO) transfektsioonireagente (Polyplus Transfection®, Prantsusmaa). Ettekasvatatud rakud loendati hemotsütoomeetriga (Hausser Scientific, USA). 24h enne transfektsiooni plaaditi välja 250 000 rakku 6-augulisele plaadile ja lisati 2 ml DMEM söödet (10% FBS, 1% pen/strep). Transfektsioon viidi läbi vastavalt tootjapoolsele protokollile. Segu inkubeeriti 10 min toatemperatuuril ja 50 µl transfektsioonisegu lisati tilkadena rakkudele. Plaati segati õrnalt 8-kujulise liigutusega ja inkubeeriti 37°C juures 24h.

2.2.5 Transfektsiooni efektiivsuse kontroll ja rakkude sorteerimine

24h möödudes kontrolliti transfektsiooni efektiivsust fluorestsentsmikroskoobiga (Mudel IX81, Olympus). Positiivse EGFP (ing k Enhanced Green Fluorescent Protein – EGFP) signaali olemasolul aspireeriti rakkudelt sööde ning rakke pesti 1 ml 1x PBS (Phosphate-Buffered Saline) lahusega. Aspireeriti PBS ja lisati 200 µl trüpsiini. Pärast rakkude tassilt lahtitulemist lisati trüpsiini inaktiveerimiseks 500 µl FBS-i sisaldavat DMEM söödet. Saadud segu pipeteeriti läbi sõelaga korgi FACS-i (ingl k Fluorescence-Activated Cell Sorting – FACS) tuubidesse, iga proov eraldi tuubi. Rakud tsentrifuugiti põhja 1000 rpm 5 min, aspireeriti sööde ning suspendeeriti 300 µl 5% FBS-i sisaldavas PBS-is. Iga proovi jaoks pandi valmis 2 FACS-i tuubi (üks EGFP positiivsete rakkude jaoks ja üks EGFP negatiivsete jaoks). EGFP positiivsetest rakkudest sorteeriti klonaalsete rakuliinide saamiseks FACS-iga üksikud raku. Üksikute kloonide jaoks lisati 96-augulisele plaadile 200 µl DMEM söödet (10% FBS, 1% pen/strep) ning igasse auku sorteeriti 1 EGFP positiivne rakk. Üksikuid kloone kasvatati 90% konfluentsuseni ja passeeriti 6-augulisele plaadile. Rakkude sorteerimine viidi läbi molekulaar- ja rakubioloogia instituudis BD FACSAria rakusorteriga (BD Bioscience).

2.2.6 DNA eraldamine

Sorteeritud rakkudest ja üksikutest kloonidest kasvatatud rakkudest eraldati genoomne DNA kasutades *DNA Clean & Concentrator™5* kitti (Zymo Research, USA). Eraldamine tehti vastavalt tootjapoolsele protokollile.

2.2.7 Illumina järgmise põlvkonna sekveneerimine

Järgmise põlvkonna sekveneerimine (NGS) viidi läbi firmas Asper Biogene, Illumina MiSeq500 masinaga. Eelnev PCR viidi läbi järgneva programmiga: algne denaturatsioon 95°C 10 min; denaturatsioon 95°C 10 sek; renaturatsioon (CRISPR-Lyl1 ja CRISPR-Lyl2 57°C, CRISPR-Tal1 61°C) 20 sek; ekstensioon 72°C 20 sek; korrata 34 tsüklit ja hoida 10°C. PCR-i segu, lõppmahus 25 µl : 2.5 µl 10x Puhver B2; 2.5 µl dNTP (2.5 mM); 2.5 µl MgCl₂ (25 mM), 1 µl F praimer (5 µM); 1 µl R praimer (5 µM); 0.2 µl Hot FirePol Taq polümeraas (5U/µl) (Solis BioDyne, Eesti), DNA 2 µl (50 ng/µl); 13.3 µl mQ. PCR-i ja sekveneerimise jaoks kasutatavad praimerid on välja toodud lisas 3. PCR-iprodukte kontrolliti geelelektroforeesiga 1.5% agarooosgeelil (70V; 1h).

2.2.8 Sanger sekveneerimine

CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 transfekteeeritud segapopulatsioone ja kloone analüüsiti Sanger sekveneerimisega, mis teostati Tartu Ülikooli Genoomika Instituudi sekveneerimise tuumiklaboris. Eelnev PCR viidi läbi järgneva programmiga: algne denaturatsioon 95°C 10 min; denaturatsioon 95°C 10 sek; renaturatsioon (CRISPR-Lyl1 ja CRISPR-Lyl2 57°C, Tal1 61°C) 20 sek; ekstensioon 72°C 20 sek; korrata 34 tsüklit ja hoida 10°C. PCR-i segu, lõppmahus 25 µl: 2.5 µl 10x Puhver B2; 2.5 µl dNTP (2.5 mM); 2.5 µl MgCl₂ (25 mM), 1 µl F praimer (5 µM); 1 µl R praimer (5 µM); 0.2 µl Hot FirePol Taq polümeraas (5U/µl), DNA 2 µl (50 ng/µl) (Solis BioDyne, Eesti); 13.3 µl mQ. PCR-i ja Sanger sekveneerimise jaoks kasutatavad praimerid on välja toodud lisas 3. PCR-iprodukte kontrolliti geelelektroforeesiga 1.5% agarooosgeelil (70V; 1h). Sobivad proovid toimetati tuumiklaborisse edasiseks analüüsiks.

2.2.9 ACT-PCR

ACT-PCR-i jaoks disainiti protokollist lähtuvalt sobilikud praimerid CRISPR-Lyl1, CRISPR-Lyl2 ja CRISPR-Tal1 (Lisa 4) jaoks nii, et *forward* praimer jääks kohakuti eeldatava kahe-ahelalise

lõikekohaga. Renaturatsiooni kriitiline temperatuur määrati kindlaks gradient PCR-iga CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 jaoks (Lisa 5). Gradient PCR programm: algne denaturatsioon 95°C 10 min; denaturatsioon 95°C 10 sek; renaturatsioon 17-nel erineval temperatuuril (55.9°C, 56.1°C, 56.8°C, 58°C, 59.4°C, 60.9°C, 62.5°C, 64.1°C, 65.6°C, 66.8°C, 67.6°C, 68°C, 69.2°C, 71°C, 72.5°C, 73.6°C, 74.1°C) 20 sek, ekstensioon 72°C 20 sek; korrata 34 tsüklit ja hoida 10°C juures. PCR-i segu, lõppmahus 25 µl: 2.5 µl 10x Puhver B2; 2.5 µl dNTP (2.5 mM); 2.5 µl MgCl₂ (25 mM), 1 µl F praimer (5 µM); 1 µl R praimer (5 µM); 0.2 µl Hot FirePol Taq polümeraas (5U/µl) (Solis BioDyne, Eesti), DNA 2 µl (50 ng/µl); 13.3 µl mQ. CRISPR-Lyl1 ja CRISPR-Lyl2 praimerite puhul jäi renaturatsiooni kriitiliseks temperatuuriks 67.2°C ja CRISPR-Tal1 jaoks 66.8°C. PCR-iprodukte kontrolliti geelelektroforeesiga 1.5% agarosgeelil 70V; 1h).

2.2.10 Bindel-PCR

Bindel-PCR-i jaoks disainiti protokollist lähtuvalt sobilikud praimerid CRISPR-Lyl1, CRISPR-Lyl2 ja CRISPR-Tal1 jaoks (Lisa 6). Esmalt määrati iga praimeripaari jaoks sobiv renaturatsiooni temperatuur gradient PCR-iga (Lisa 7A). Gradient PCR programm: algne denaturatsioon 95°C 10 min; denaturatsioon 95°C 10 sek; renaturatsioon 12-nel erineval temperatuuril (55.9°C, 56.1°C, 56.8°C, 58°C, 59.4°C, 60.9°C, 62.5°C, 64.1°C, 65.6°C, 66.8°C, 67.6°C, 68°C) 20 sek; ekstensioon 72°C 20 sek; korrata 34 tsüklit (Nested-PCR 30 tsüklit) ja hoida 10°C juures. PCR-i segu, lõppmahus 25 µl: 2.5 µl 10x Puhver B; 2.5 µl dNTP (2.5 mM); 2.5 µl MgCl₂ (25 mM), 1 µl F praimer (5 µM) (Tal_san_F; Lyl_san_F); 1 µl R praimer (5 µM) (Tal_san_R; Lyl_san_R); 0.2 µl Hot FirePol Taq polümeraas (5U/µl) (Solis BioDyne, Eesti), DNA 2 µl (50 ng/µl); 13.3 µl mQ. Sama protokollil alusel määrati nested-PCR-i praimerite sobivad sulamistemperatuurid. Järgnevalt määrati sobiva MgCl₂ kontsentratsioon (Lisa 7B), mille juures metsik-tüüpi DNA puhul amplifitseeritakse PCR produkt. PCR-i segu, lõppmahus 25 µl: 2.5 µl 10x Puhver B2; 2.5 µl dNTP (2.5 mM); x µl MgCl₂ (2.5 mM; 2 mM; 1.5 mM; 1 mM; 0.85 mM; 0.7 mM), 1 µl F praimer (5 µM); 1 µl R praimer (5 µM); 0.2 µl Hot FirePol Taq polümeraas (5U/µl) (Solis BioDyne, Eesti), DNA (50 ng/µl); x µl mQ (maht viidi 25 µl-ni). PCR-iprodukte kontrolliti geelelektroforeesiga 1.5% agarosgeelil (70V; 1h).

2.2.11 ORNi-PCR

Iga proovi jaoks disainiti 3 erineva pikkusega oligonukleotiidi (Lisa 8), mille 3' ots on amino blokeeritud. ORNi PCR-i jaoks kasutatud praimerid on välja toodud lisas 3. PCR-i programm: algne denaturatsioon 95°C 10 min; denaturatsioon 95°C 10 sek; renaturatsioon (CRISPR-Lyl1 ja CRISPR-Lyl2 57°C; CRISPR-Tal1 61°C) 20 sek; ekstensioon 72°C 20 sek; korrata 34 tsüklit ja hoida 10°C juures. PCR-i segu, lõppmahus 50 µl: 2,5 µl F praimer (5 µM); 2,5 µl R praimer (5 µM); 0,5 µl DNA (50 ng/µl); 14,5 µl mQ; 25 µl NEBNext® High-Fidelity 2X PCR Master Mix (New England BioLabs, USA), 5 µl oligonukleotiide (10 µM). PCR-i produkte kontrolliti geelelektroforeesiga 1.5% agarosgeelil (70V; 1h). Testitud oligonukleotiidide seast valiti välja kõige efektiivsemad (Lisa 9)

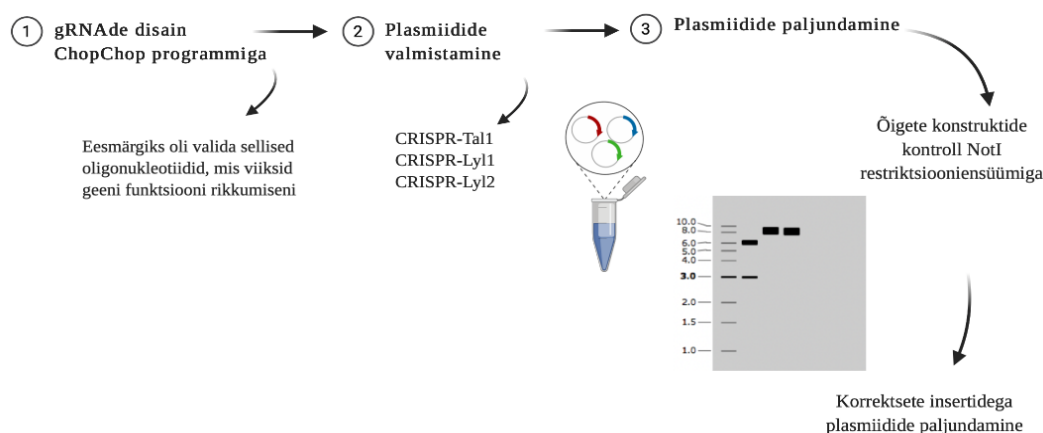
2.2.12 Illumina NGS ja Sanger sekveneerimise andmete analüüs

Illumina MiSeq500 paired-end fastq failid joondati mm10 (*Mus musculus*) referentsgenoomile, kasutades Bowtie2 (Langmead et al., 2009) programmi. Bowtie2 parameetrid olid järgmised: -N 1; -k 2. Kasutades samtools 0.1.19 (Li et al., 2009) versiooni indekseeriti ja sorteeriti bam failid. Saadud bam formaadis andmeid visualiseeriti esmalt tarkvaraprogrammiga IGB (Integrated Genome Browser) (Freese et al., 2016). Andmete analüüsimiseks kasutati statistikatarkvara R lisapaketti CrispRVariant (Lindsay et al., 2016). Sanger sekveneerimisest saadud .ab1 formaadis toorandmeid visualiseeriti esmalt BioEdit tarkvaraga. Edasi analüüsiti andmeid erinevate programmidega nagu Synthego ICE (Synthego Performance Analysis, ICE Analysis. 2019. v2.0. Synthego; [2020]) ja R põhise paketiga CrispRVariant (Lindsay et al., 2016).

2.3 Tulemused

2.3.1 Plasmiidide valmistamine

Edukaks CRISPR/*Cas9* *knock-out* eksperimendiks on vajalik nii *Cas9* valgu kui ka gRNA ekspressioon rakkudes. gRNA järjestuse ja *Cas9* ekspresseerimiseks NIH3T6 rakkudes kasutati pX-458 *NotI* plasmidi (Lisa 2), kuhu oli sisestatud sihtmärkjärjestusele disainitud oligonukleotiidid (Joonis 8). Selle plasmidi transfekteeerimise efektiivsust on tänu EGFP ekspressioonile võimalik lihtsasti kontrollida. Antud plasmiidile on lisaks olemasolevale *NotI* äratundmisjärjestusele lisatud ka teine *NotI* restriktsooniensüümi lõikekoht, mis kloneerimise käigus eemaldatakse ja seetõttu on antud restriktsooniensüümiga võimalik kontrollida, kas kloneerimise käigus on plasmidi integreerunud soovitud oligonukleotiidne järjestus. Korrektsete insertide olemasolu plasmidis kinnitati esmalt *NotI* restriktsooniensüümiga (Lisa 10) ning seejärel Sanger sekveneerimisega (Lisa 11).



Joonis 8. Praktilise osa I etapi katseskeem plasmidi valmistamisest. Esimese etapina sai disainitud oligonukleotiidid, mis paikneksid selliselt, et saaks rikutud geeni funktsioon. Oligonukleotiidid kloneeriti pX-458 *NotI* plasmidi ja selliselt valmistati 3 erinevat plasmidi: CRISPR-Ta11, CRISPR-Lyl1 ja CRISPR-Lyl2. Plasmide paljundati kompetentsetes rakkudes, misjärel puhastati plasmiidid ning kontrolliti *NotI* restriktsooniensüümiga.

2.3.2 NIH3T6 rakuliini transfekterimine CRISPR-Lyl1, CRISPR-Lyl2 ja CRISPR-Tal1 plasmiidiga

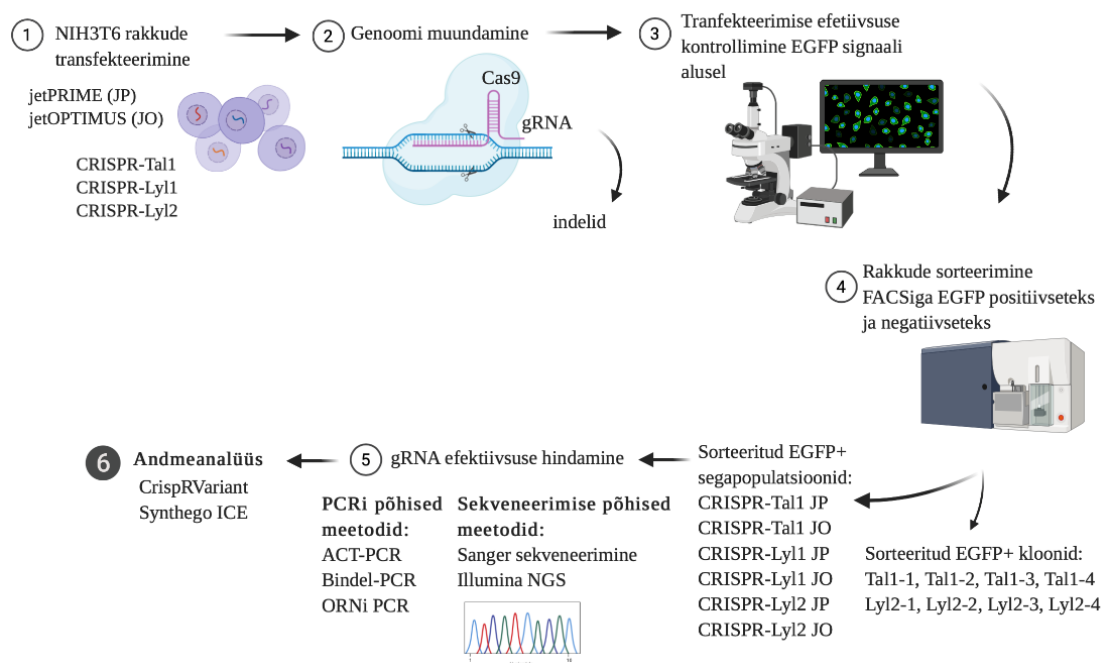
NIH3T6 rakke transfekteriti kolme plasmiidiga, millest kaks olid suunatud Lyl1 geeni ja üks Tal1 geeni vastu (CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2). 24h möödudes kontrolliti transfekterimise efektiivsust EGFP signaali alusel fluorestsentsmikroskoobiga (Joonis 9). Joonisel 10A on visuaalsel hinnangul näha, et kõikide plasmiididega - CRISPR-Lyl1, CRISPR-Lyl2 ja CRISPR-Tal1 on transfekterimine õnnestunud. Samas EGFP signaali olemasolu üksi ei võimalda hinnata disainitud gRNA-de efektiivsust - kas gRNA on edukalt *Cas9* juhtinud soovitud sihtmärgini ja kas sihtmärkjärjestuses on *Cas9* poolt tekitatud kahe-ahelalise katke parandamisel tekkinud mutatsioone. Selleks, et testida erinevate meetoditega gRNA efektiivsust, sai FACS-iga sorteeritud EGFP positiivsed segapopulatsioonid (CRISPR-Tal1 JP, CRISPR-Tal1 JO, CRISPR-Lyl1 JP, CRISPR-Lyl1 JO, CRISPR-Lyl2 JP, CRISPR-Lyl2 JO) ning EGFP negatiivsed rakupopulatsioonid (CRISPR-Tal1 JP-, CRISPR-Lyl1 JP-, CRISPR-Lyl2 JP-) (Tabel 3, Lisa 12). Neid rakupopulatsioone analüüsiti edasi töös välja toodud kolme PCR-i põhise meetodiga: ACT-PCR, Bindel-PCR ja ORNi PCR-iga ning kahe sekveneerimise meetodiga: Sanger sekveneerimise ja Illumina järgmise põlvkonna sekveneerimisega.

Tabel 3. FACS-iga sorteeritud rakupopulatsioonid

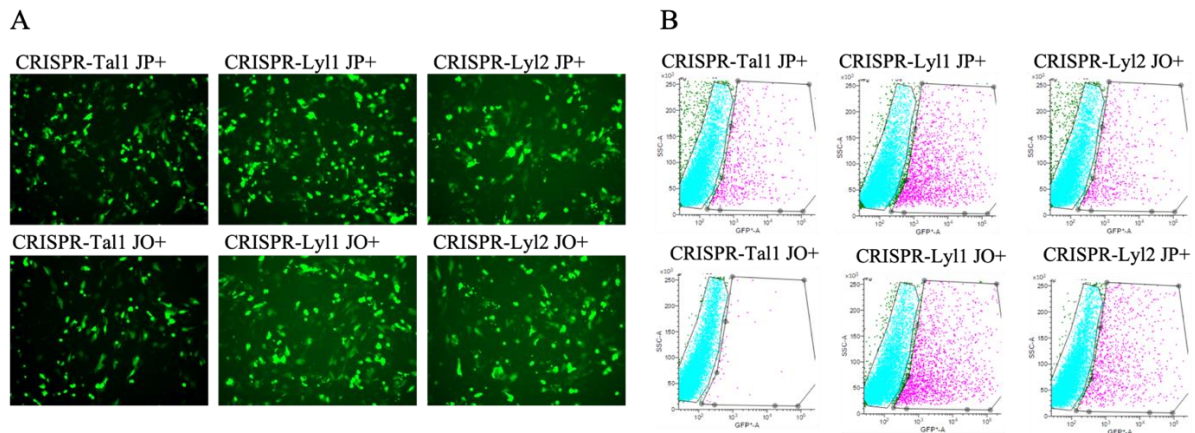
Proovid	EGFP+ segapopulatsioonid	EGFP- segapopulatsioonid	Kloonid
	CRISPR-Tal1 JP	CRISPR-Tal1 JP-	Tal1-1
	CRISPR-Lyl1 JP	CRISPR-Lyl1 JP-	Tal1-2
	CRISPR-Lyl2 JP	CRISPR-Lyl2 JP-	Tal1-3
	CRISPR-Tal1 JO		Tal1-4
	CRISPR-Lyl1 JO		Lyl2-1
	CRISPR-Lyl2 JO		Lyl2-2
			Lyl2-3
			Lyl2-4

Viie korduskatse põhjal oli jetPPRIME®-iga transfekterimise efektiivsus FACS-iga analüüsimisel CRISPR-Tal1 puhul 8%, CRISPR-Lyl1 20% ning CRISPR-Lyl2 transfekterimise efektiivsus oli keskmiselt 24%. FACS-iga sorteerides saadi kolme plasmidi puhul keskmiselt 27076 EGFP positiivset rakku ja 146712 EGFP negatiivset rakku (Lisa 13). Meetodite efektiivsuse

ja tulemuste kontrollimiseks sai välja sorteeritud ka EGFP positiivsetest rakkudest üksikud kloonid 96-augulisele plaadile. Selliselt sisaldaks iga kлонаalne populatsioon eeldatavasti sama mutatsiooni ning proovides puuduksid muundamata järjestused, mistõttu neid proove oleks võimalik kasutada positiivse kontrollina PCR-i põhistes meetodites. Sellisel juhul ei oleks proovides ülekaalus metsik-tüüpi järjestused, mille tõttu alahinnatakse indeleid sisaldavate järjestuste osakaalu.



Joonis 9. Praktilise osa II etapi katseskeem. Eelnevas etapis valmistatud kolme plasmiidiga-CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 transfecteeriti NIH3T6 rakke. Transfecteerimiseks kasutati kahte transfectsioonireagenti: jetPRIME® ja jetOPTIMUS® ning 24h möödudes kontrolliti transfecteerimise efektiivsust fluorestsentsmikroskoobiga. Positiivse signaali olemasolul sorteeriti rakud FACS-iga EGFP positiivseteks segapopulatsioonideks (CRISPR-Tal1 JP, CRISPR-Tal1 JO, CRISPR-Lyl1 JP, CRISPR-Lyl1 JO, CRISPR-Lyl2 JP ja CRISPR-Lyl2 JO) ja EGFP negatiivseteks (CRISPR-Tal1 JP-, CRISPR-Lyl1 JP-, CRISPR-Lyl2 JP-). Lisaks sai sorteeritud EGFP positiivsetest rakkudest üksikud kloonid Tal1-1, Tal1-2, Tal1-3, Tal1-4 ja CRISPR-Lyl2 jaoks Lyl2-1, Lyl2-2, Lyl2-3, Lyl2-4. Muundatud rakupopulatsioonide analüüsimiseks kasutati nii PCR-i, kui ka sekvenerimise põhiseid meetodeid.



Joonis 10. Transfektiooni efektiivsuse kontroll ja rakkude sorteerimine FACS-iga. **A** CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 plasmidid transfektiooni kontroll fluorestsentsmikroskoobiga. **B** EGFP signaali alusel sorteeritud EGFP positiivsed ja EGFP negatiivsed rakupopulatsioonid, mida analüüsiti töös väljatoodud 3 PCR-i põhise meetodiga: ACT-PCR, Bindel-PCR ja ORNi ja kahe sekveneerimise meetodiga: Sanger sekveneerimise ja Illumina järgmise põlvkonna sekveneerimisega.

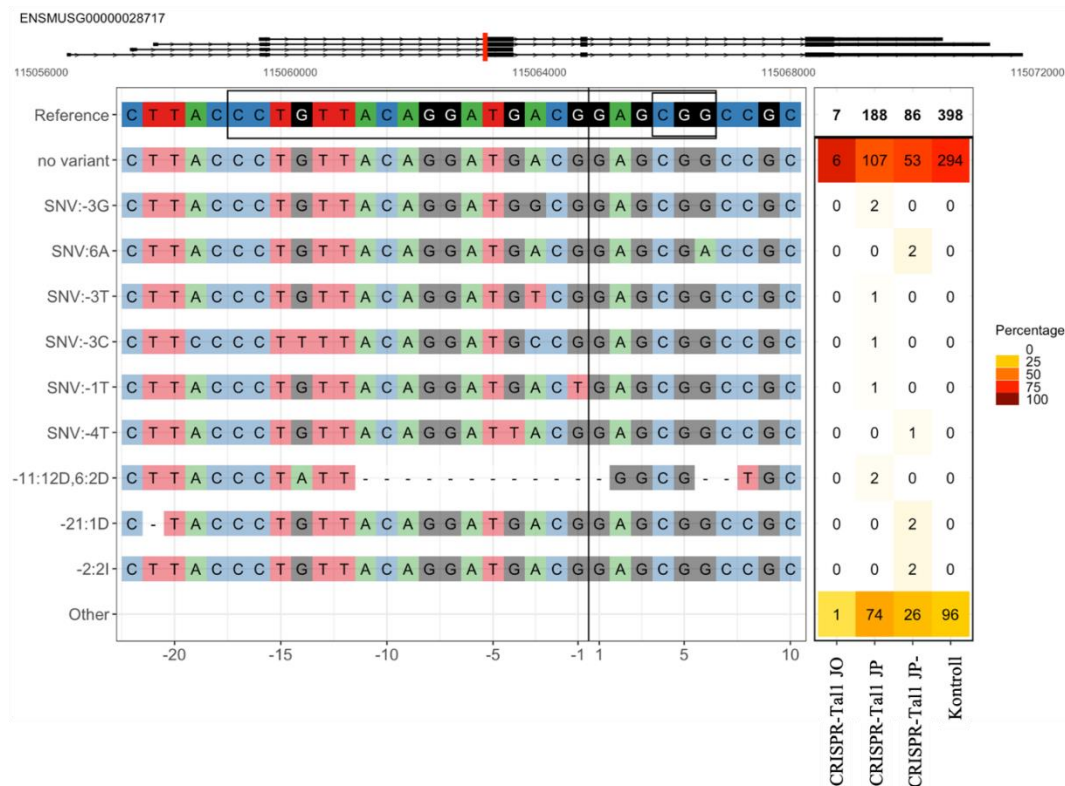
2.3.3 CRISPR-Tal1, CRISPR-Lyl1, CRISPR-Lyl2 segapopulatsioonide ja klonaalsete rakkude NGS tulemused

Selleks, et saada ülevaadet tekkinude indelitest, said CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 transfekteeritud segapopulatsioonid (tähistatud JP ja JO) ja klonalsed rakupopulatsioonid sekveneeritud Illumina *MiSeq*-ga ja Sanger sekveneerimisega.

Illumina NGS tulemusi analüüsiti R põhise lisapaketi CrispRVariantiga, mille väljundiks on tabel, kus on näha järjestustes esinevate indelite sagedused protsentides ja indelite positsioonid järjestuses (Joonis 11). CrispRVarianti sisendandmeteks on referentsgenoomile joondatud bam failid, milles võrreldakse gRNA genoomsete koordinaatide põhjal muutusi analüüsitavates failides kasutades referentsina metsik-tüüpi järjestusi.

NGS tulemuste analüüsil CrispRVariantiga oli näha, et CRISPR-Tal1 JP muundamise efektiivsus oli 41,18%, kuid CRISPR-Tal1 JO puhul oli indeleid sisaldavad lugemeid väga vähe, mistõttu ei saa antud proovi puhul hinnata gRNA efektiivsust adekvaatselt hinnata. Joonisel 11 on välja toodud CRISPR-Tal1 segapopulatsioonide CrispRVariant analüüsi tulemused, milles on esitatud 10 kõige sagedasemat indeli tüüpi. CRISPR-Tal1 JP ja CRISPR-Tal1 JO analüüsitud segapopulatsioonides moodustasid ~3,7% indelitest 1 nukleotiidilised muutused, mis asuvad 3 nukleotiidi ülesvoolu eeldatavast *Cas9* lõikepositsioonist (nullpunktiks on 3 nukleotiidi ülesvoolu PAM-ist ehk eeldatav

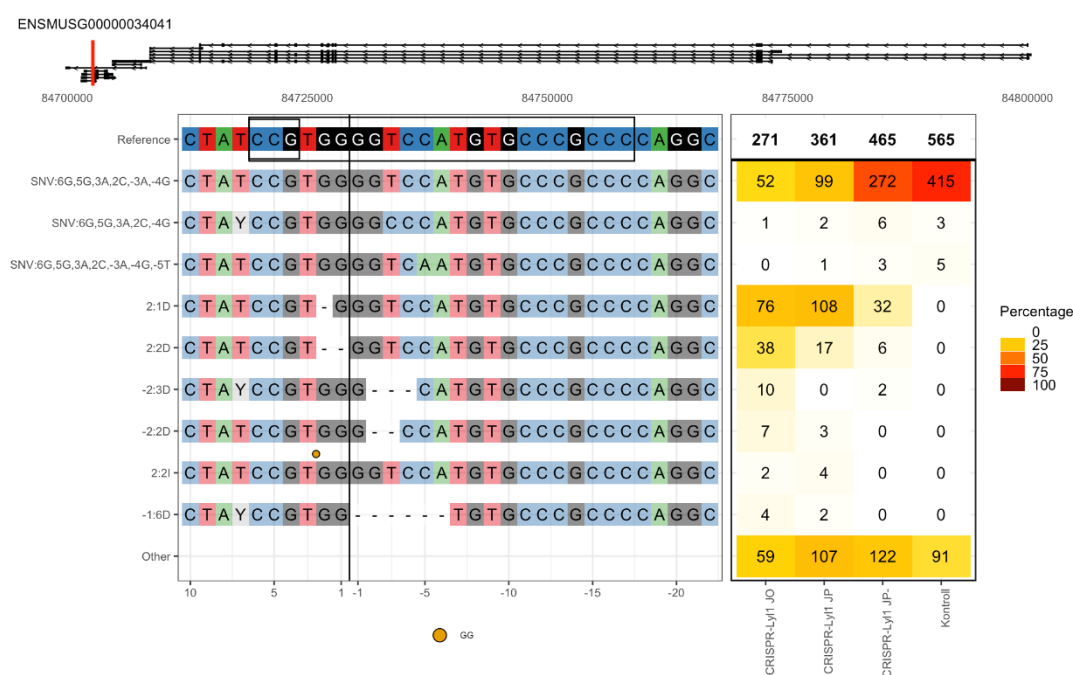
Cas9 lõikepositsioon). Kõige ulatuslikum indel CRISPR-Tal1 JP segapopulatsioonis on 11 nukleotiidiline deletsioon. Antud andmete analüüsist on näha, et ka EGFP negatiivseks populatsiooniks sorteeritud CRISPR-Tal1 JP- proovis esinevad indelid (Joonis 11). Antud juhul võib olla põhjuseks mitmeid: rakkude tihedus on sorteerimisel olnud liiga suur või sorteerimise värvad on määratud liiga lähedale negatiivsele fraktsioonile. Kuna rakud sorteeriti 24h peale transfekkeerimist on võimalik, et osades rakkudes ei olnud sorteerimise hetkel EGFP signaal veel detekteeritav.



Joonis 11. CRISPR-Tal1 segapopulatsioonide Illumina NGS andmete analüüsi tulemused CrispRVariantiga. Joonisel on esitatud 10 kõige sagedasemat indeli tüüpi, mis esinevad CRISPR-Tal1 segapopulatsioonide proovides. Joonise ülemises servas on märgitud gRNA positsioon geenil. Musta kastiga on tähistatud gRNA järjestus ning 3 nukleotiidiline PAM. Musta läbiva joonega on tähistatud eeldatav *Cas9* lõikepositsioon, mis asub 3 nukleotiidi ülesvoolu PAMist. Vasakul on välja toodud indelide tüübid ja positsioonid järjestuses. Paremal on *heatmap*, millel on esitatud indeleid sisaldavate lugemite arv. CRISPR-Tal1 segapopulatsiooni proovides olid kõige sagedasemateks variatsioonideks 1 nukleotiidilised asendused, lisaks esines ka kuni 11 nukleotiidilisi deletsioone.

CRISPR-Lyl1 puhul oli sihtmärkjärjestuse muundamise efektiivsus keskmiselt 74%. Nendest 38% moodustasid ka kirjanduse alusel kõige sagedasemad 1 ja 2 nukleotiidilised deletsioonid (Joonis 12). Samuti leidis antud proovide hulgas ka erineva pikkusega 9-13 nukleotiidilisi deletsioone ning

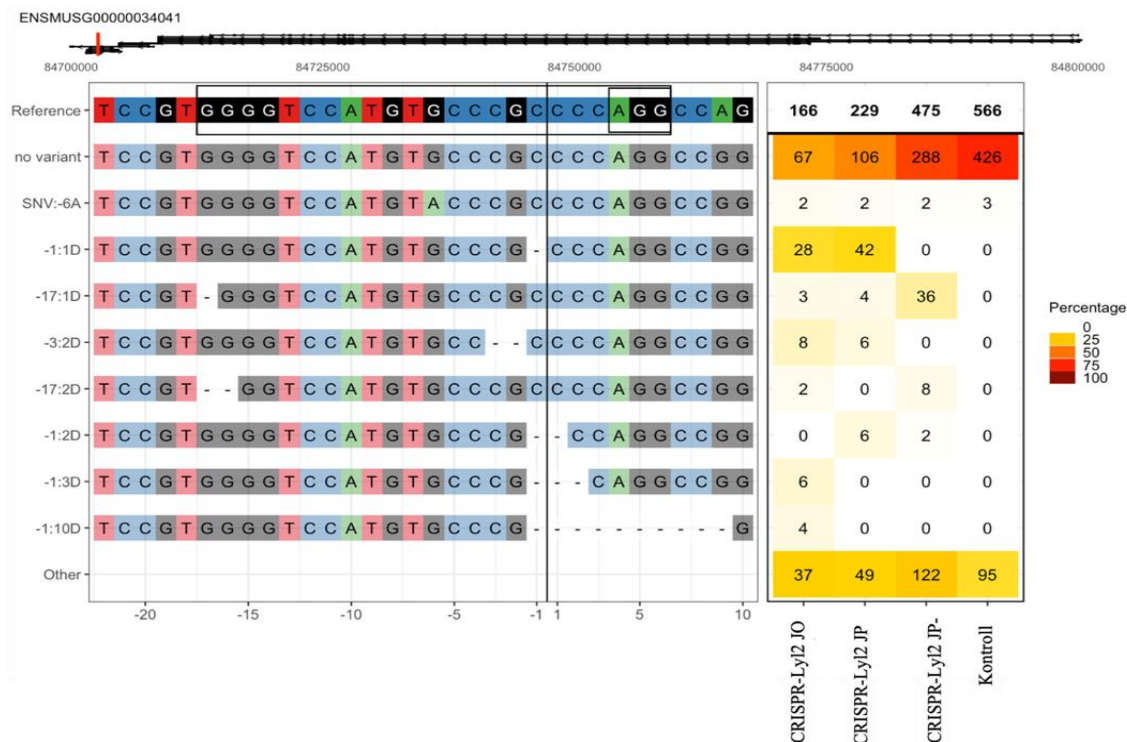
1-9 nukleotiidiliseid insertioone. Joonisel 12 on näha, et EGFP alusel negatiivseks sorteeritud rakupopulatsioonides CRISPR-Lyl1 JP- esineb 1 ja 2 nukleotiidilisi deletsioone. Tõenäoliselt võib ka siin olla põhjuseks mittepiisava tugevusega EGFP signaal või FACS-iga sorteerimisel tekivad vead. Samuti leidub ka kontroll proovis ühe nukleotiidilisi muutusi. Arvatavasti võivad sellised muutused tekkida kas juba esialgse PCR-i või siis sekveneerimise käigus. Ka kirjanduses on näidatud, et 39% ühe nukleotiidilistest asendustest on seletatavad polümeraasi töös tekkivate vigadega. Samas on ka NGS-i veamääraks kirjanduse põhjal antud $>0,1\%$ ning kombinatsioonis PCR-i käigus tekkivate vigadega, on loogiline, et sellised ühe nukleotiidilised asendused võivad järjestustes esineda.



Joonis 12. CRISPR-Lyl1 segapopulatsioonide Illumina NGS andmete analüüsi tulemused CrispRVariantiga. Joonisel on esitatud 10 kõige sagedasemat indeli tüüpi, mis esinevad CRISPR-Lyl1 segapopulatsioonide proovides. Joonise ülemises servas on märgitud gRNA positsioon geenis. Musta kastiga on tähistatud gRNA järjestus ning 3 nukleotiidiline PAM. Musta läbiva joonega on tähistatud eeldatav *Cas9* lõikepositsioon, mis asub 3 nukleotiidi ülesvoolu PAMist. Vasakul on välja toodud indelide tüübid ja positsioonid järjestuses. Paremal on *heatmap*, millel on esitatud indeleid sisaldavate lugemite arv. Kollase ringiga on tähistatud insertiooni positsioon ning nukleotiidne järjestus. CRISPR-Lyl1 segapopulatsiooni proovides olid kõige sagedasemad 1 ja 2 nukleotiidilised indelid, lisaks esines ka kuni 8 nukleotiidilisi deletsioone.

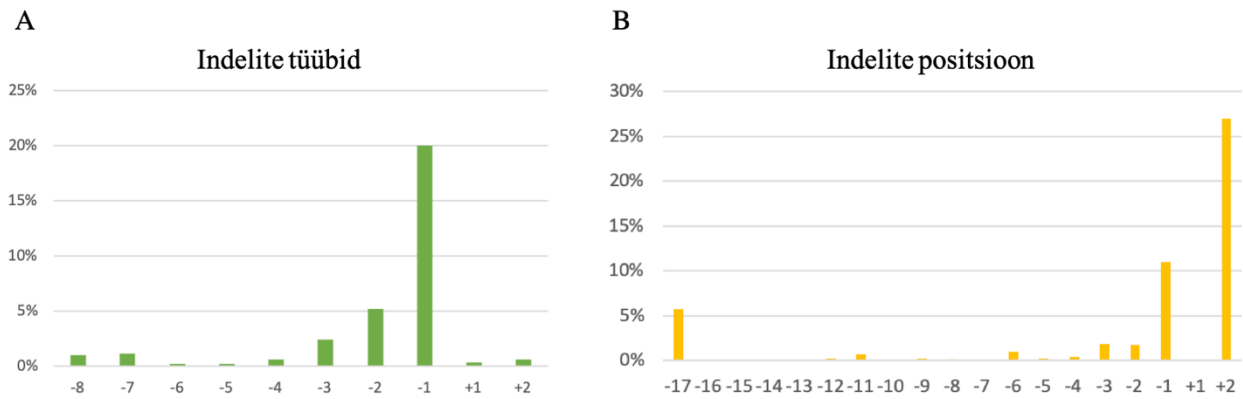
CRISPR-Lyl2 puhul oli muundamise efektiivsus keskmiselt 55%. Sarnaselt CRISPR-Lyl1 ja CRISPR-Tal1 proovidele moodustas ka CRISPR-Lyl2 muundatud järjestustes kõige suurema osa

(~28%) 1 ja 2 nukleotiidilised deletsioonid (Joonis 13). Samuti esines ka CRISPR-Lyl2 proovides ulatuslikke (9-13 nukleotiidseid) deletsioone ning mõned üksikud insertioonid, kus järjestusse oli lisatud 1-4 nukleotiidi. CRISPR-Lyl2 puhul 33% indelitest asub sellises positsioonis, mis viivad geeni funktsiooni rikkumiseni.



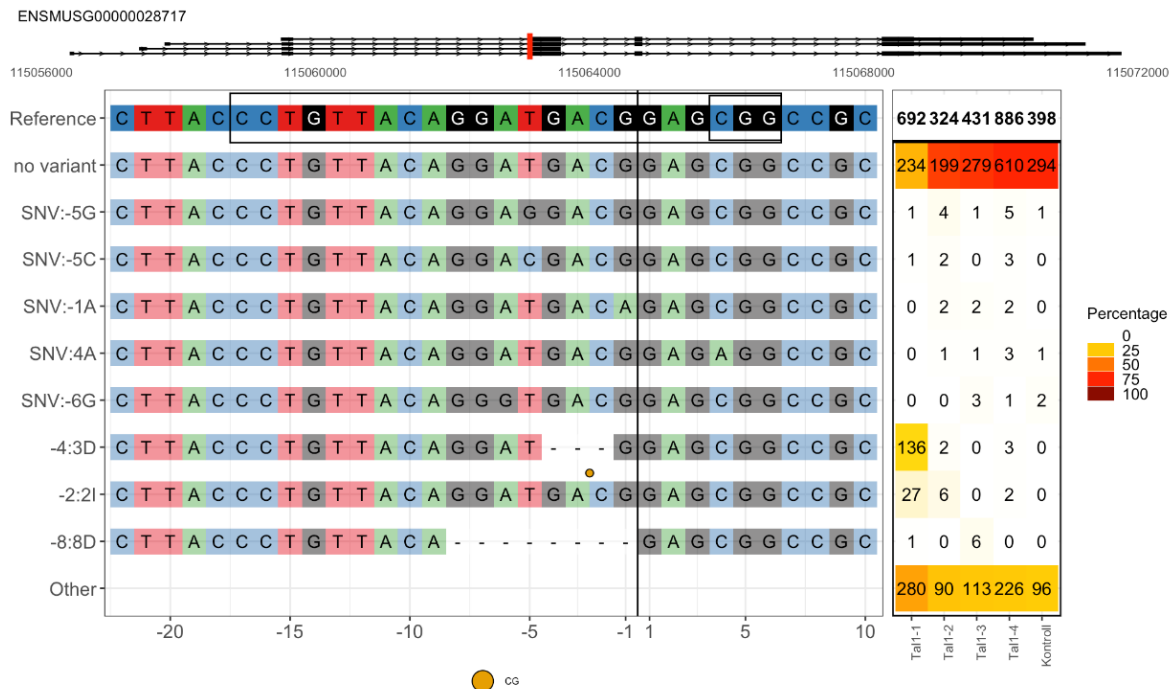
Joonis 13. CRISPR-Lyl2 segapopulatsioonide Illumina NGS andmete analüüsi tulemused CrispRVariantiga. Joonisel on esitatud 10 kõige sagedasemat indeli tüüpi, mis esinevad CRISPR-Lyl2 segapopulatsioonide proovides. Joonise ülemises servas on märgitud gRNA positsioon geenis. Musta kastiga on tähistatud gRNA järjestus ning 3 nukleotiidiline PAM. Musta läbiva joonega on tähistatud eeldatav *Cas9* lõikepositsioon, mis asub 3 nukleotiidi ülesvoolu PAMist. Vasakul on välja toodud indelide tüübid ja positsioonid järjestuses. Paremal on *heatmap*, millel on esitatud indeleid sisaldavate lugemite arv. CRISPR-Lyl2 segapopulatsiooni proovides oli kõige sagedasemaks indeliteks 1 ja 2 nukleotiidilised indelid, lisaks esines ka kuni 10 nukleotiidilisi deletsioone.

Joonisel 14 on kokkuvõtvalt esitatud CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 segapopulatsioonides esinevad kõige sagedasemad indelid ning nende positsioonid järjestustes. Antud juhul on saadud tulemused sarnased kirjanduses esitatud tulemustega. Kõige sagedasemateks indeliteks on 1-3 nukleotiidilised deletsioonid, mis paiknevad *Cas9* lõikepositsiooni läheduses.



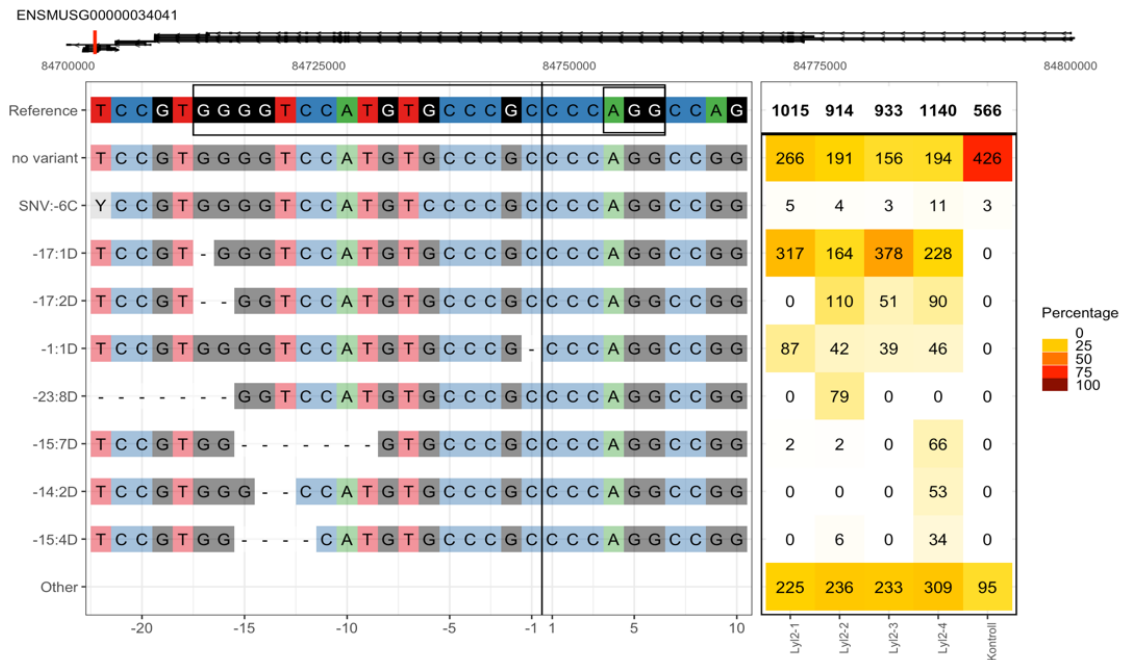
Joonis 14. Segapopulatsioonides esinevad kõige sagedasemad indelite tüübid ja positsioonid. CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 segapopulatsioonides esinevate indelite tüübid ja positsioonid NGS analüüsi põhjal. A Joonisel on esitatud insertioonide ja deletsioonide pikkused (-1 tähistab 1 nukleotiidilist deletsiooni; +1 tähistab 1 nukleotiidilist insertiooni) ning nende sagedused protsentides. Kõige suurema osakaalu analüüsitud proovides moodustasid 1, 2 ja 3 nukleotiidilised deletsioonid. **B** Esitatud on analüüsitud proovides esinevate indelite positsioonid (0 positsiooniks on 3 nukleotiidi ülesvoolu PAMist).

Analüüsides CRISPR-Tal1 ja CRISPR-Lyl2 klonaalseid rakupopulatsioone selgus, et kõik populatsioonid sisaldavad kombinatsioone erinevatest indelitest. CRISPR-Tal1 ja CRISPR-Lyl2 klonaalsete populatsioonides olid kõige sagedasemateks mutatsioonideks 1-3 nukleotiidilised deletsioonid (Joonis 15, Joonis 16). Üldiselt leidis kõikides CRISPR-Tal1 klonaalsetes proovides erineva sagedusega erinevat tüüpi indeleid, mille hulka kuulusid ulatuslikud (6-11 nukleotiidilised) deletsioonid, 2 nukleotiidilised insertioonid ja ühe nukleotiidilised muutused. Vaatamata sellele, et kõik antud proovid olid FACS-iga sorteeritud üksikud kloonid, esineb kõikides proovides erinevat tüüpi indeleid. Põhjuseks võib olla see, et antud rakkudes oli CRISPR/*Cas9* aktiivne veel pärast sorteerimist ning uutes rakkudes toimetas CRISPR/*Cas9* edasi, mistõttu tekkis populatsioonis kombinatsioon erinevatest indelitest. Samas moodustab suurema osakaalu antud proovides metsik-tüüpi järjestused, mis tähendab, et CRISPR/*Cas9* efektiivsus on antud proovides jäänud väga madalaks. Klonaalsetest proovidest sisaldas kõige suurema osakaaluga indeleid vaid Tal1-1, teistes klonaalsetes proovides on indelite osakaal väga väike ning pigem esineb nendes üksikuid ühe nukleotiidilisi muutusi.



Joonis 15. CRISPR-Tal1 klonaalsetes rakupopulatsioonides esinevad indelid. Joonisel on esitatud 10 kõige sagedasemat indeli tüüpi, mis esinevad CRISPR-Tal1 klonaalsetes proovides (Tail-1, Tail-2, Tail-3, Tail-4). Joonise ülemises servas on märgitud gRNA positsioon geenis. Musta kastiga on tähistatud gRNA järjestus ning PAM. Musta läbiva joonega on tähistatud eeldatav *Cas9* lõikepositsioon, mis asub 3 nukleotiidi ülesvoolu PAMist. Vasakul on välja toodud indelide tüübid ja positsioonid järjestuses. Paremal on *heatmap*, millel on esitatud indeleid sisaldavate lugemite arv. Kollase ringiga on tähistatud kahe nukleotiidilise insertiooni positsioon ning järjestus. Kõige sagedasemateks indelideks CRISPR-Tal1 klonaalsetes populatsioonides on 3 nukleotiidilised deletsioonid ning 2 nukleotiidilised insertioonid.

CRISPR-Lyl2 klonaalsetes proovides moodustasid 40% kogu järjestustest 1-2 nukleotiidilised deletsioonid (Joonis 16). Võrreldes CRISPR-Tal1 proovidega esineb CRISPR-Lyl2 klonaalsetes proovides vähem 1 nukleotiidilisi muutusi, kuid suurema sagedusega leidub ulatuslike, kuni 22 nukleotiidilisi deletsioone.



Joonis 16. CRISPR-Ly12 klonaalsetes rakupopulatsioonides esinevad indelid. Joonisel on esitatud 10 kõige sagedasemat indeli tüüpi, mis esinevad CRISPR-Ly12 klonaalsetes proovides (Ly12-1, Ly12-2, Ly12-3, Ly12-4). Joonise ülemises servas on märgitud gRNA positsioon geenis. Musta kastiga on tähistatud gRNA järjestus ning PAM. Musta läbiva joonega on tähistatud eeldatav *Cas9* lõikepositsioon, mis asub 3 nukleotiidi ülesvoolu PAMist. Vasakul on välja toodud indelide tüübid ja positsioonid järjestuses. Paremal on heatmap, millel on esitatud indeleid sisaldavate lugemite arv. Kõige sagedasemateks indelideks CRISPR-Ly12 klonaalsetes populatsioonides on 1 ja 2 nukleotiidilised deletsioonid.

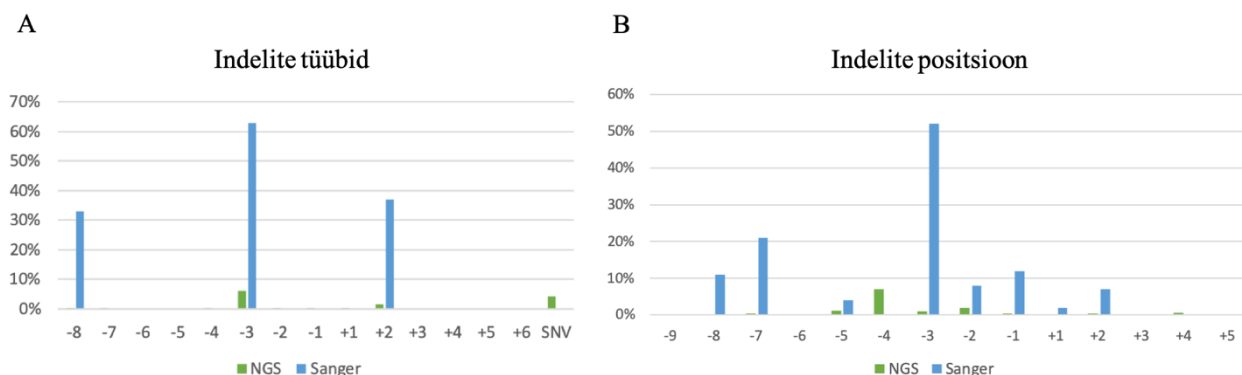
2.3.4 CRISPR-Tal1, CRISPR-Ly11, CRISPR-Ly12 segapopulatsioonide ja klonaalsete rakkude Sanger sekveneerimise tulemused

Muundatud järjestustes esinevate indelide tuvastamiseks on kasutatud ka traditsioonilist Sanger sekveneerimist. Meetodi toimimise põhimõtte tõttu on seda meetodit kasutatud peamiselt üksikute rakkude analüüsiks. Samas on loodud mitmeid programme nagu Synthego ICE, mis võimaldab analüüsida Sanger sekveneerimisega ka segapopulatsiooni proove. NGS-iga analüüsitud sega- ja klonaalseid populatsioone analüüsiti ka Sanger sekveneerimisega, et näha, kas selliselt on võimalik kindlaks määrata segapopulatsioonides esinevad indelide tüübid ja kui sarnased on saadud tulemused NGS-i tulemustega. Analüüsiks kasutatud Synthego ICE sai valitud programmi lihtsa kasutajaliidese ning tulemuste hea visuaalse esitamise tõttu. Synthego ICE sisendfailideks on Sanger sekveneerimisest saadud .ab1 failid, millest üks on metsik-tüüpi, muundamata järjestus ning teine muundatud järjestus. gRNA järjestuse järgi pannakse paika

eeldatav *Cas9* lõikepositsioon ning muundatud järjestust võrreldakse kasutaja poolt antud metsik-tüüpi järjestusega. Synthego ICE-i puhul esitatakse tulemused nii kromatogrammna, kui ka nukleotiidses järjestuses, kuhu on märgitud tekkinud indelite positsioon, sagedus ja suurus. Synthego ICE antud juhul ei analüüsinud CRISPR-Tal1 segapopulatsioone tundmatu vea 334 pärast. Tõenäoliselt mõjutab Sanger sekveneerimisest saadud andmete analüüsi ka indelite osakaal - juhul kui ühe vahemikku jääb palju erinevaid indeleid ei pruugi programmid kromatogrammis esinevaid indeleid eristada. Samas suutis Synthego ICE edukalt määrata CRISPR-Tal1 klonaalsetes populatsioonides esinevad indelite tüübid. CRISPR-Tal1 klonaalsete proovide puhul oli kõige sagedasemateks indelideks Synthego ICE tulemuste põhjal 3 nukleotiidilised deletsioonid (Joonis 17, Joonis 18). Kuigi Tal1-4 rakud olid üles kasvatatud FACS-iga sorteeritud ühest EGFP positiivsest rakust, puudusid Synthego ICE analüüsi tulemuste põhjal antud proovis sihtmärkjärjestuses indelid. Selline olukord võib tekkida, kui algselt sortiti EGFP negatiivne rakk, gRNA *Cas9* kompleks ei töötanud efektiivselt või kaheaahelaline katke parandati ilma vigadeta.



Joonis 17. CRISPR-Tal1 klonaalsete populatsioonide Sanger sekveneerimise andmete analüüs Synthego ICE-iga. Vasakul on tähistatud indelite tüübid ja suurus ning indelite esinemissagedus protsentides. Paremal on näidatud indelite positsioonid *Cas9* eeldatava lõikepositsiooni suhtes.

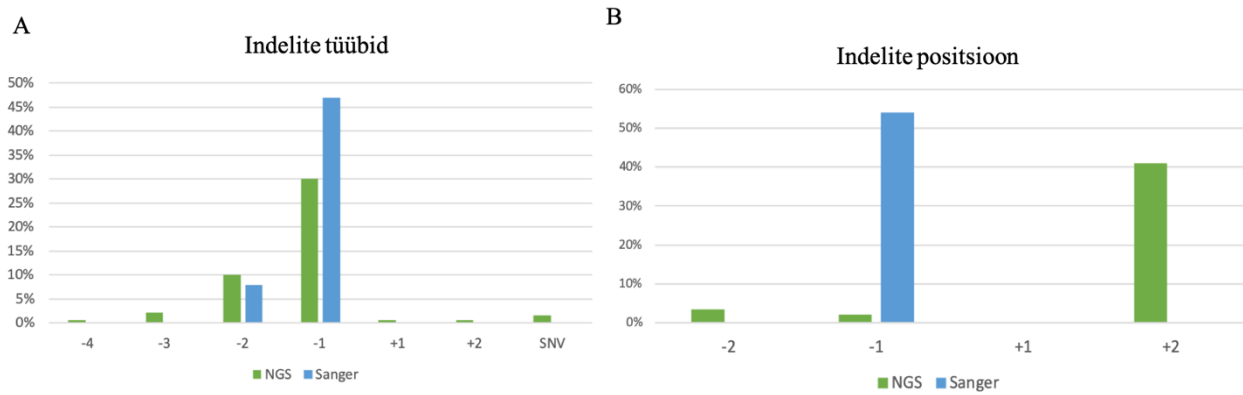


Joonis 18. CRISPR-Tal1 klonaalsetes proovides esinevate indelite positsioonid ja tüübid NGS ja Sanger sekveneerimise alusel. **A** NGS ja Sanger sekveneerimisega tuvastatud indelite tüübid. SNV tähistab ühe nukleotiidilisi asendusi. Miinusega on esitatud deletsioonide pikkused ning plussiga insertiooni pikkused. **B** Joonisel on esitatud indelite positsioonilise jaotuvuse erinevused NGS ja Sanger sekveneerimise vahel. 0 positsiooniks on 3 nukleotiidi ülesvoolu PAMist.

Analüüsid Synthego ICE-iga CRISPR-Lyl1 segapopulatsioone, määras antud programm kõige sagedasemateks indeliteks 1 nukleotiidilised deletsioonid (Joonis 19, Joonis 20A). Antud juhul on NGS-i ja Sanger sekveneerimise tulemused erinevad, eriti indelite positsiooni määramisel (Joonis 20B). NGS-i tulemuste põhjal asus 38% indelitest positsioonis +2 kuid Synthego ICE on määranud enamuse indelid positsiooni -1.

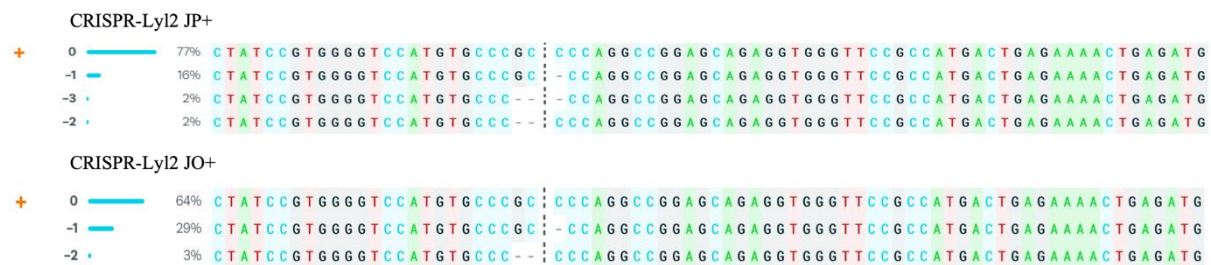


Joonis 19. CRISPR-Lyl1 segapopulatsioonide Sanger sekveneerimise andmete analüüs Synthego ICE-iga. Vasakul on tähistatud indelite tüübid ja suurus ning indelite esinemissagedus protsentides. Paremal on näidatud indelite positsioonid *Cas9* eeldatava lõikepositsiooni suhtes. Antud proovide puhul on kõige sagedasemateks indeliteks 1 nukleotiidilised deletsioonid, kuid esineb ka insertioone ning kuni 17 nukleotiidilisi deletsioone.

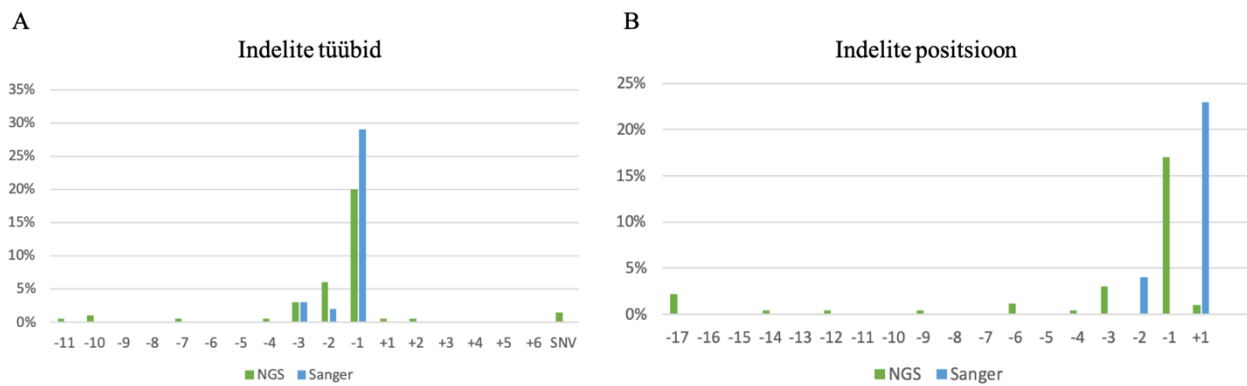


Joonis 20. CRISPR-Lyl1 segapopulatsioonide indelite positsioonid ja tüübid NGS ja Sanger sekveneerimise alusel. **A** NGS ja Sanger sekveneerimisega tuvastatud indelite tüübid. SNV tähistab ühe nukleotiidilisi asendusi. Miinusega on esitatud deletsioonide pikkused ning plussiga insertiooni pikkused. **B** Joonisel on esitatud indelite positsiooniline jaotuvus NGS ja Sanger sekveneerimise tulemuste põhjal. 0 positsiooniks on 3 nukleotiidi ülesvoolu PAMist.

CRISPR-Lyl2 Sanger sekveneerimisest saadud andmete analüüsil jättis Synthego ICE analüüsist välja CRISPR-Lyl2 klonalsed rakupopulatsioonid ning andis tulemused vaid CRISPR-Lyl2 segapopulatsioonide kohta. Antud programm andis põhjuseks vea failide formaadis. CRISPR-Lyl2 JP ja CRISPR-Lyl2 JO analüüsitud proovides olid kõige sagedasemateks indeliteks samuti 1 ja 2 nukleotiidilised deletsioonid, mis asusid 2 ja 3 nukleotiidi ülesvoolu PAMist (Joonis 21). Antud juhul olid NGSi ja Sanger sekveneerimise tulemused omavahel üsna sarnased indelite tüüpide määramisel (Joonis 22A). Mõlema analüüsi põhjal oli kõige sagedasemaks indeliks 1 nukleotiidilised deletsioonid. Samas erinevad üksteisest selgelt indelite positsioon, NGS-i põhjal asusid enamus indelid positsioonis -1 ja -3, kuid Sanger sekveerimisest saadud andmete põhjal on Synthego ICE määranud enamus indelid positsiooni -1 (Joonis 22B).



Joonis 21. CRISPR-Lyl2 segapopulatsioonide Sanger sekveneerimise andmete analüüs Synthego ICE-iga. Vasakul on tähistatud indelite tüübid ja suurus ning indelite esinemissagedus protsentides. Paremal on näidatud indelite positsioonid *Cas9* eeldatava lõikepositsiooni suhtes. Antud proovide puhul on kõige sagedasemateks indeliteks 1 ja 2 nukleotiidilised deletsioonid.



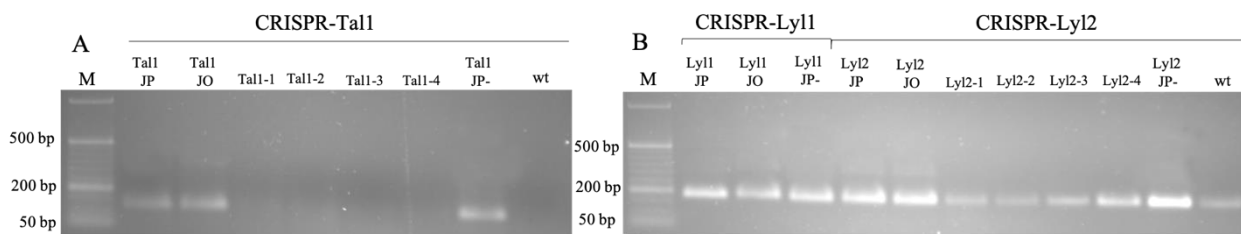
Joonis 22. CRISPR-Lyl2 segapopulatsioonide indelite positsioonid ja tüübid NGS ja Sanger sekveneerimise alusel. A NGS ja Sanger sekveneerimisega tuvastatud indelite tüübid. SNV tähistab ühe nukleotiidilisi muutusi. Miinusega on esitatud deletsioonide pikkused ning plussiga insertiooni pikkused. B Joonisel on esitatud indelite positsiooniline jaotuvuse erinevused NGS ja Sanger sekveneerimisega saadud tulemuste põhjal. 0 positsiooniks on 3 nukleotiidi ülesvoolu PAMist.

Vaadates kahe sekveneerimismeetodiga saadud tulemusi on selgelt näha, et NGS on võimeline oluliselt paremini tuvastama erinevaid indelid ning nende positsioone (Joonis 23). Kahe meetodi omavahelised tulemused ei korreleeru omavahel väga hästi.

2.3.5 ACT-PCRi tulemused

ACT-PCR põhineb alleel-spetsiifilisel PCR-il, kus kasutatakse mutatsioonitundlike primereid kombineerituna praimerite kriitilise seondumistemperatuuriga. Indeleid sisaldavad proovid on agarosgeelil negatiivsed, kuna indelide esinemise korral ei toimu praimer 3' otsa seondumist sihtmärkjärjestusega. Joonisel 23 on välja toodud ühed representatiivsed ACT-PCR-i tulemused, millest on näha, et CRISPR-Tal1 proovidest on geelelektroforeesiga kontrollides positiivse tulemuse andnud CRISPR-Tal1 JP, CRISPR-Tal1 JO ja CRISPR-Tal1 JP- segapopulatsioonide proovid. NGS-i tulemuste põhjal on näha, et antud proovid sisaldavad pigem metsik-tüüpi järjestusi ning indeleid sisaldavaid lugemeid on väga vähe, mistõttu on need proovid visualiseeritavad geelil. Samas pole PCR-i amplifikatsioon pole õnnestunud klonaalsete rakupopulatsioonidega Tal1-1, Tal1-2, Tal1-3 ja Tal1-4, mis antud analüüsi järgi viitab sellele, et need proovid sisaldavad sihtmärkjärjestuses indeleid. NGS-i tulemuste põhjal sisaldas vaid Tal1-1 proov selliseid indeleid, mille puhul oleks tõesti praimer 3' otsa seondumine häiritud. Ülejäänud klonaalsetes proovides esineb indeleid väga vähe ning ülekaalus on siiski metsik-tüüpi järjestused, mille puhul ei oleks

praimeri 3' otsa seondumine häiritud. Samas pole antud juhul PCR-i amplifikatsiooni õnnestunud ka kontroll prooviga (wt). Kuna CRISPR-Tal1 nelja korduskatse puhul ei töötanud korduvalt ka kontrollid, pole antud juhul saadud tulemused usaldusväärsed.

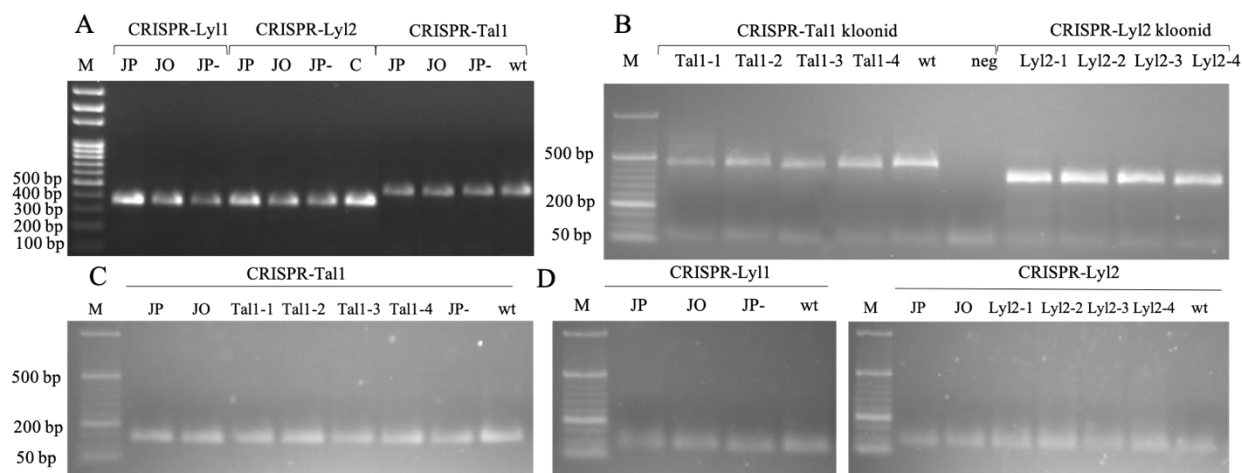


Joonis 23. CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 ACT-PCRi tulemused. Joonisel on näha representatiivsed geelipildid neljast korduskatsest. **A** CRISPR-Tal1 segapopulatsioonide ja klonaalsete rakupopulatsioonide tulemused. Positiivse tulemuse on andnud CRISPR-Tal1 JP, CRISPR-Tal1 JO ja CRISPR-Tal1 JP-. PCR-i amplifikatsioon pole õnnestunud CRISPR-Tal1 klonaalsete proovidega Tal1-1, Tal1-2, Tal1-3 ja Tal1-4. Samuti ei töötanud ootuspäraselt kontroll proov, mis peaks olema visualiseeritav geelil. **B** CRISPR-Lyl1 ja CRISPR-Lyl2 segapopulatsioonide ja CRISPR-Lyl2 klonaalsete populatsioonide tulemused. Positiivsed tulemused on andnud kõik analüüsitavad proovid, nii segapopulatsioonid (CRISPR-Lyl1 JP, CRISPR-Lyl1 JO, CRISPR-Lyl1 P-, CRISPR-Lyl2 JP, CRISPR-Lyl2 JO, CRISPR-Lyl2 JP-, Lyl2-1, Lyl2-2, Lyl2-3, Lyl2-4). Samuti on töötanud kontroll proov, kuid oodatust madalama efektiivsusega.

CRISPR-Lyl1 ja CRISPR-Lyl2 puhul on kõik analüüsitud proovid andnud eeldatava pikkusega PCR-i produkti (~190 bp), mis antud analüüsi järgi tähendab, et nendes järjestustes on suurema osakaaluga sellised muutused, mille puhul PCR-i amplifikatsioon ei ole takistatud. CRISPR-Lyl1 NGS andmete analüüsimisel CrispRVariantiga oli näha, et enamus (47%) indelitest jäi praimeri 3' otsa, seega peaksid CRISPR-Lyl1 proovid olema negatiivsed. Sarnaselt on ka CRISPR-Lyl2 JP ja JO segapopulatsiooni proovidega toimunud PCR-i amplifikatsioon, antud proovid sisaldasid NGS andmete analüüsi põhjal vaid 26% selliseid indeleid, mille puhul oleks ACT-PCR F praimeri 3' otsa seondumine takistatud. Seega on ülekaalus sellised järjestused, mille puhul toimub F praimeri seondumine sihtmärkjärjestusega ning tegelikult ei ole võimalik selliselt anda hinnangut gRNA efektiivsuse kohta.

2.3.6 Bindel-PCRi tulemused

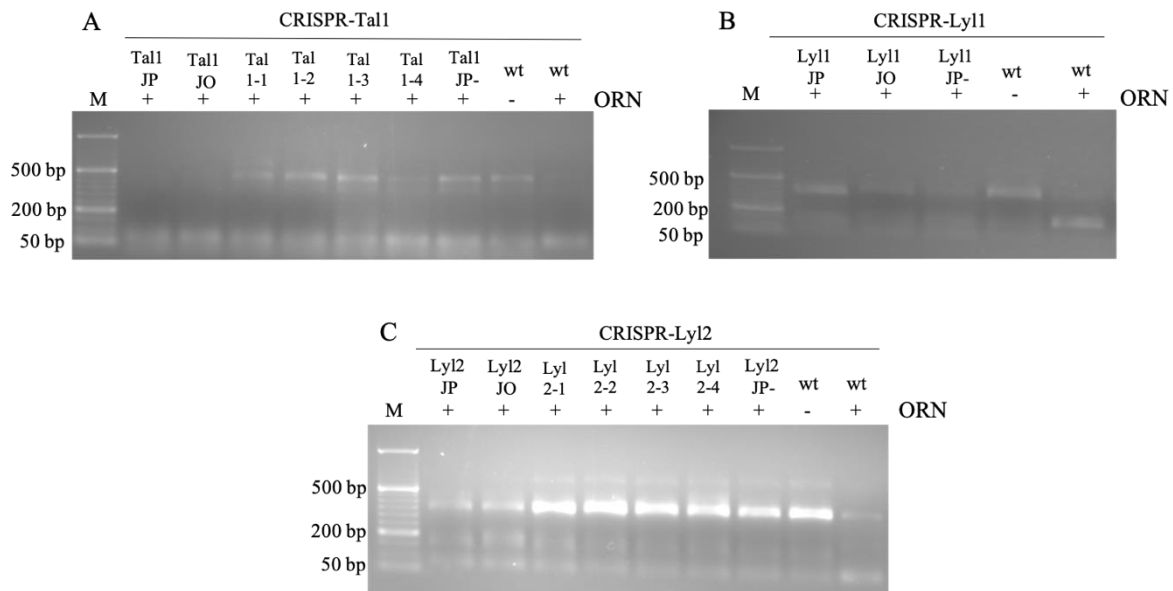
Bindel-PCR põhineb alleel-spetsiifilisel PCR-il, mis on kombineeritud $MgCl_2$ kontsentratsiooni ja *nested*-PCR-iga. Esimene PCR viidi läbi standardtingimustel, kasutades primereid, mis on disainitud selliselt, et saadud ampliconid sisaldavad sihtmärkjärjestust. Joonisel 24A ja 24B on esitatud esimese PCR-i tulemused, kus on näha, et kõikide proovidega on saadud oodatud suurusega PCR-i produkt (CRISPR-Tal1 ~432 bp, CRISPR-Lyl1, CRISPR-Lyl2 ~357 bp). *Nested*-PCR-is kasutati madalamat, 0.85 mM $MgCl_2$ kontsentratsiooni (Lisa 7B) ja mutatsiooni tundlike primereid, mis viiksid PCR-i ebaõnnestumiseni indelide esinemisel. *Nested*-PCRi tulemustest on näha, et kõikide proovidega on saadud õige suurusega PCR-i ampliconid (Joonis 24C, 24D). Samas peaksid indeleid sisaldavad proovid meetodi artikli kohaselt olema geelelektrofooresiga kontrollides negatiivsed. Selle meetodi puhul alahinnatakse kõige sagedasemaid, 1 ja 2 nukleotiidilisi indeleid, kui need jäävad primer 5' otsa või vahetult väljapoole primeri sihtmärkjärjestusest. Ka selle meetodi puhul saab määravaks metsik-tüüpi järjestuste osakaal, sest selliste järjestuste puhul toimub tõenäoliselt edukas PCR-i amplifikatsioon.



Joonis 24. Bindel-PCR tulemused. **A** Bindel-PCR esimese PCR-i tulemused, kõikide analüüsitud proovidega on saadud oodatud suurusega PCR-i ampliconid. CRISPR-Tal1 puhul 432 bp ja CRISPR-Lyl1 ja CRISPR-Lyl2 357 bp. **B** CRISPR-Tal1 (Tal1-1 – Tal1-4) ja CRISPR-Lyl2 (Lyl2-1 – Lyl2-4) kloonide PCR-i amplifikatsioon. **C** CRISPR-Tal1 segapopulatsioonide ja klonaalsete populatsioonide nested-PCR 0.85 mM $MgCl_2$ kontsentratsiooniga. **D** CRISPR-Lyl1 ja CRISPR-Lyl2 segapopulatsioonide ja klonaalsete populatsioonid nested-PCR 0.85 mM $MgCl_2$ kontsentratsiooniga.

2.3.7 ORNi PCR

Kolmas meetod - ORNi PCR põhineb oligonukleotiididel, mis blokeerivad muundatud proovides metsik-tüüpi järjestuselt toimuva PCR-i. CRISPR-Tal1 proovide puhul kasutati PCR-i reaktsioonis 26 nukleotiidilist oligonukleotiidi ning CRISPR-Lyl1 ja CRISPR-Lyl2 proovide jaoks 24 nukleotiidilist oligonukleotiidi. Joonisel 25A on näha, et eeldatavasti, kui meetod on töötanud ootuspäraselt, on proovides blokeeritud PCR metsik-tüüpi järjestustelt ning PCR on toimunud indeleid sisaldavatel järjestustelt. CRISPR-Tal1 puhul on PCR blokeeritud proovides CRISPR-Tal1 JP ja CRISPR-Tal JO, see on ka NGS-i tulemuste põhjal oodatav. Kuna antud proovides oli indeleid sisaldavaid järjestusi väga vähe, on tõesti eeldatav, et need proovid on geelil negatiivsed. Kuid antud juhul põhineb see hinnang PCR-i bändi intensiivsusel ning eeldusel, et PCR-is on tõesti blokeeritud metsik-tüüpi järjestused. Osaliselt on PCR-i amplifikatsioon toimunud proovides Tal1-1, Tal1-2, Tal1-3, CRISPR-Tal1 JP- ning vähesel määral Tal1-4 proovis. Samas CRISPR-Tal1 klonaalsetes proovid sisaldavad pea kõik proovid sihtmärkjärjestuses indeleid, kuid ka nende osakaal oli väiksem kui muundamata järjestuste osakaal. Kõige suurema osakaaluga sisaldas indeleid Tal1-1 klonaalne proov, millest võiks eeldada, et PCR toimub efektiivsemalt, kuid ka selle proovi puhul on PCR-i amplifikatsioon jäänud üsna madalaks. Tal1-2, Tal1-3 ja Tal1-4 proovides esines rohkem ühe nukleotiidilisi muutusi ning metsik-tüüpi järjestusi, mistõttu eeldaks, et PCR-i amplifikatsioon oleks nende proovide puhul takistatud, sest sobivatel tingimustel võib ka blokeeriv oligonukleotiid tolereerida järjestuse keskel olevaid 1 nukleotiidilisi muutusi. Sellisel juhul blokeeritaks ära ka need järjestused, milles esinevad kõige sagedasemad 1 nukleotiidilised indelid. CRISPR-Lyl1 analüüsitud proovides on kõikides proovides toimunud erineva efektiivsusega PCR-i amplifikatsioon (Joonis 25B). Kontroll proovidega on töötanud PCR ootuspäraselt: kontrollid kuhu on lisatud blokeeriv oligonukleotiid (Joonis 25A wt+, joonis 25B wt+ ja joonis 25C wt+) on üldiselt metsik-tüüpi järjestustelt PCR blokeeritud. Samuti on ootuspärased tulemused ka kontrollidega milles puuduvad blokeerivad oligonukleotiidid (Joonis 25A wt-, joonis 25B wt- ja joonis 25C wt-) - toimunud on edukas PCR-i amplifikatsioon. Samas on CRISPR-Lyl1 kontrollis näha ebaspetsiifilise pikkusega amplikoni, kus peaks olema blokeeritud PCR-i amplifikatsioon. Kuna sihtmärkjärjestusest sõltuvalt on disainitud oligonukleotiidid GC rikkad, võib sellisel juhul tekkida sekundaarstruktuure ja oligonukleotiidide dimeer, mis võivad anda geelelektroforeesiga kontrollides erineva suurusegaprodukte.



Joonis 25. CRISPR-Tal1, CRISPR-Lyl1 ja CRISPR-Lyl2 ORNi PCR-i tulemused. Joonisel on näidatud representatiivsed geelipildid viiest korduskatsest. ORN tähistab blokeerivat oligonukleotiidi ning wt tähistab metsik-tüüpi kontrolli proovi. **A** CRISPR-Tal1 segapopulatsioonide ja klonaalsete populatsioonide ORNi PCR-i tulemused. **B** CRISPR-Lyl1 ORNi PCR-i tulemused. **C** CRISPR-Lyl2 segapopulatsioonide ja klonaalsete proovide ORNi PCR-i tulemused

CRISPR-Lyl2 proovide puhul on kõikides proovides nähtavad PCR-i bändid (Joonis 25C), mis on ka oodatav. NGS andmete analüüsist selgus, et CRISPR-Lyl2 proovid sisaldasid ~50% juhul selliseid indeleide, mille puhul võiks olla oligonukleotiidi seondumine häiritud ning nendelt järjestustelt saaks toimuda PCR-i amplifikatsioon. Siiski oli ka nendes proovides (CRISPR-Lyl2 JP ja CRISPR-Lyl2 JO) selliseid indeleide, mis jäid väljapoole oligonukleotiidi sihtmärkjärjestusest. Selle tõttu oleks siiski PCR blokeeritud ka indeleide sisaldavate järjestuste puhul.

2.4 Arutelu

CRISPR/*Cas9* genoomi muundamise eksperimendi üheks oluliseks sammuks on gRNA efektiivsuse hindamine. Tänapäeval on olemas lai valik erinevaid meetodeid, mida võib selleks kasutada. Need erinevad nii kulude, aja, kui ka saadud andmete osas. Lisaks vajavad erinevad meetodid erinevaid seadmeid, mis võivad ulatuda kõige tavalisemast PCR-i termotsüklerist ja geelelektroforeesi süsteemidest moodsaate suure jõudlusega sekveneerimisseadmeteni. Kahjuks ei ole üheski lihtsas ja odavas meetodis kombineeritud kõige kriitilisemad tingimused: täpsus, kiirus ja kulutõhusus.

NGS andmete analüüsil oli selgelt näha, et suurema osakaalu indelitest moodustavad lühikesed, 1-3 nukleotiidilised deletsioonid (CRISPR-Lyl2 28%, CRISPR-Tal1 3,7%, CRISPR-Lyl1 40%), mis on ka kooskõlas kirjanduses esitatud tulemustega (Allen *et al.*, 2019; Chakrabarti *et al.*, 2019; van Overbeek *et al.*, 2016). Samuti esines analüüsitud proovides ulatuslike (5-22 nukleotiidilisi) deletsioone, mis jäid nii ülesvoolu kui allavoolu PAM-ist, kuid nende osakaal on üldiselt väiksem (Taheri-Ghahfarokhi *et al.*, 2018). Pikemad deletsioonid võivad olla põhjustatud mikrohomoloogia poolt suunatud otste parandamisest (Taheri-Ghahfarokhi *et al.*, 2018; van Overbeek *et al.*, 2016). Arvatavasti tekivad pikemad deletsioonid seetõttu, et mikrohomoloogia poolt suunatud otste parandamisel eemaldatakse kahe homoloogilise piirkonna (2-20 nukleotiidi) vahel olev järjestus ning seejärel ühendatakse homoloogilised otsad (Ata *et al.*, 2018; Sfeir ja Symington, 2015). Kuigi ka antud töös analüüsitud proovides esines pikemaid deletsioone, ei ole võimalik öelda, kas need on tekkinud mikrohomoloogia poolt suunatud parandamise tagajärjel. Kuna mikrohomoloogia poolt suunatud parandamine on aeglasem ja toimub pikema aja vältel, on loogiline, et segapopulatsioonidel puhul antud parandismehhanismile viitavaid tunnuseid ei leitud, sest antud populatsioonid sorteeriti ja eraldati DNA 24h peale transfekterimist. Kirjanduses tuuakse üha rohkem välja ka seda, et kõige sagedasemate indelite positsioon ei ole juhuslik, vaid ennustatav *Cas9* lõikepositsiooni ümbritseva sihtmärkjärjestuse järgi. Kuna antud töös testiti vaid kolme gRNA-d, ei saanud antuid väiteid kinnitada, sest järjestustes esinevate indelite põhjal ei tekkinud kindlat mustrit, mille põhjal oleks võimalik näidata, et üks või teine nukleotiid oleks olnud eelistatud.

Kuna Sanger sekveneerimine on üsna odav ja lihtne meetod, sai olemasolevad proovid analüüsida ka selle meetodiga. Selliselt sai võrrelda, kui sarnased on Sanger sekveneerimise tulemused NGS-i tulemustega. Kahe sekveneerimise meetodi võrdlemisel oli näha, et Sanger sekveneerimise tulemused erinevad mõnevõrra NGS-i tulemustest. Olulisemad erinevused tulid välja indelite

positsiooni määramisel ja indelite sageduses, mida oli selgelt näha CRISPR-Ly11 ja Tal1-4 proovide puhul. Sangeri sekveneerimisest saadud failide analüüsimisel on oluline failide kvaliteet ja indelite sagedus. Lisaks võivad nõrga signaali ja taustamüra tõttu tuvastamata jääda harvemini esinevad indelid. Kuna ka järjestuse vahemik, mida analüüsitakse on piiratud kindlate parameetrite poolest, ei pruugi Sangeri sekveneerimine tuvastada pikemaid deletsioone.

PCR-i põhiste meetodite puhul saab määravaks teguriks gRNA efektiivsus ja metsik-tüüpi järjestuste osakaal proovides. Kuna gRNAd töötavad erineva efektiivsusega (Doench *et al.*, 2014; Mandal *et al.*, 2014; Xu *et al.*, 2015), sisaldavad muundatud järjestused ka metsik-tüüpi järjestusi. Kuna töös testitud kolmest PCR-i põhisest meetodist kaks esimest, ACT- ja Bindel-PCR töötavad alleel-spetsiifilise PCR-i põhimõttel, on üsna selge, et nende meetoditega alahinnatakse indelite tegeliku osakaalu proovides. Nende meetodite õnnestumine on otseselt sõltuv indelite positsioonist, disainitud praimeritest ja metsik-tüüpi järjestuste osakaalust muundatud proovides. Antud kahe meetodi puhul on oluline, et disainitud praimer 3' ots jääks positsiooni kus esineb kõige suurema sagedusega indeleid. Kuigi indelite positsioon on ennustatav PAM motiivi järgi, võivad tekkida indelid ka väljapoole sihtmärkjärjestust või jääda sellisesse positsiooni, millega ei ole mõjutatud praimer 3' otsa seondumine (Green *et al.*, 2015). Seega saab pea alati toimuda metsik-tüüpi järjestuselt PCR-i amplifikatsioon ja need tulemused on visualiseeritavad geelil. Nii ACT-PCR, kui ka Bindel-PCR-i puhul on tulemused tõlgendatavad vaid PCR-i amplikonide intensiivsuse põhjal. Selliselt ei ole aga katse tulemused ühtselt hinnatavad ja tulemused jäävad umbmääraseks ning tegelikult ei saa nende põhjal hinnata gRNA efektiivsust. Kuna PCR-i amplifikatsiooni mõjutavad peale praimerite seondumise ka teised tegurid, mille puhul võib olla PCR-i amplifikatsioon inhibeeritud ning geelelektroforeesiga kontrollides negatiivne. Ka antud töö puhul oli näha, et isegi kui PCR-i tingimused on optimeeritud, võib juhtuda, et korduvalt ei tööta metsik-tüüpi kontroll, mille põhjal sai PCR-i tingimused optimeeritud. Mõlemate meetodite, ACT-PCR ja Bindel-PCR-i, tulemuste analüüs põhineb PCR-i bändi puudumisel ja seetõttu on raske öelda, kas PCR-i amplifikatsioon oli takistatud tõesti järjestustes esinevatest indelist, või on põhjuseks muud tegurid, mis mõjutavad PCR-i. Sobivates tingimustes piisavate PCR-i tsüklite korral saab alati lõpuks geelelektroforeesiga kontrollides positiivse tulemuse. Kui ka korduskatsed omavahel erinevad ning ei ole võimalik katse tulemusi korrata, kaotavad need meetodid usaldusväärsuse. Kui hakata aga muutma katsetingimusi, et saada meetodi artiklile sarnaseid tulemusi, siis võib saada samuti kallutatud tulemused. Väga hästi tulid välja sellised probleemid Bindel- ja ACT-PCR-i puhul, kus mõlema meetodi tulemused on mõjutatavad juba väikeste

eksperimentaalsete muutustega. ACT-PCR-i puhul oli määravaks teguriks praimerite sulamistemperatuur ja Bindel-PCR-i puhul $MgCl_2$ kontsentratsioon. Selgelt oli tulemustest näha, et antud meetodid ei sobi gRNA efektiivsuse hindamiseks, kuna neis meetodites on PCR-i tingimused viidud kriitilisele piirile (kõige kõrgem praimerite sulamistemperatuur ja kõige madalam $MgCl_2$ kontsentratsioon), mille tõttu ei ole võimalik tulemusi korrektselt hinnata.

Võrreldes kahe eelnimetatud PCR-i meetodiga, andis kõige paremini tõlgendatavad tulemused ORNi PCR. ORNi meetodi artiklis kasutatakse oligoribonukleotiide, mis peaksid võimaldama stabiilsemat paardumist sihtmärgiga ning ära hoidma seda, et neid praimeritena kasutaks. Artikkel väidab, et ORN-id on odavad, kuid tellides näiteks Metabionist, maksab üks oligoribonukleotiid ~95€. Kui võtta arvesse soovitusi, et iga proovi jaoks võiks testida vähemalt kolme oligoribonukleotiidi, tuleb see üsna kallis, ~285€ ühe proovi kohta ning meetod ei ole enam odav. Sellepärast sai antud töö jaoks tellitud oligonukleotiidid, mis on 3' otsast amino modifitseeritud ja mida ei ole seetõttu võimalik PCR-i praimerina kasutada. Selle meetodi põhimõtte järgi peaksid disainitud oligonukleotiidid blokeerima PCR-i proovides olevatelt metsik-tüüpi DNA järjestustelt ning amplifikatsioon saab sellisel juhul toimuda vaid indeleid sisaldavatelt järjestustelt. Kuna antud meetodi puhul kasutatakse kahte kontrolli, millest ühele on lisatud blokeerivat oligonukleotiidi ja teisel mitte, on nende põhjal oluliselt lihtsam tulemusi tõlgendada. Ka sellel meetodil on omad miinused, mille tõttu tulemuste tõlgendamisel võidakse alahinnata populatsioonis esinevaid indeleid. Töös kasutatud proovide analüüsist oli näha, et kõigi kolme kasutatud gRNA puhul oli kõige sagedasemateks mutatsioonideks 1 ja 2 nukleotiidilised indelid. Juhul kui indelite positsioon jääb oligonukleotiidi 5' otsa võidakse sobivates tingimustes ka indeleid sisaldavad järjestused ära blokeerida ning seeläbi ei toimuks PCR-i ka indeleid sisaldavatelt järjestustelt. Lisaks on alati võimalus, et indelid võivad jääda ka oligonukleotiidi sihtmärkjärjestusest väljapoole mistõttu samuti alahinnataks indelite osakaalu, sest ka muundatud järjestused on samuti blokeeritud. Antud meetodi puhul on väga oluline ka kasutatav polümeraas ning oligonukleotiidi seondumistemperatuur. Meetodi jaoks sobivad eelkõige polümeraasid, millel puudub 5' -> 3' eksonukleaasne aktiivsus, kuna vastasel juhul blokeeriv oligonukleotiid lihtsalt lagundatakse. Nagu tuli välja ka meetodi artiklis, võib madalama seondumistemperatuuriga oligonukleotiid elongatsiooni käigus sihtmärkjärjestuse küljest lahti tulla. Selliselt saab algselt vale-positiivsed tulemused ning peab uuesti optimeerima sobivaid tingimusi PCR-i jaoks, mis omakorda võib anda kallutatud tulemusi.

Kuna testitud kolme meetodiga ei olnud võimalik saada samu tulemusi, mis oli saadud originaal meetodite artiklites, on tegelikult väga raske nende järgi hinnata, kas rakupopulatsioonid sisaldavad sihtmärkjärjestuses indeleid. Kui tegemist oleks olnud pimekatsega, ei oleks ühegi PCR-i põhise meetodi põhjal võimalik õigesti hinnata gRNA tegelikku efektiivsust. Kõiki testitud PCR-i põhiseid meetodeid peab iga sihtmärgi jaoks eraldi optimeerima, mistõttu antud meetodid lähevad üsna ajakulukaks, eriti kui on vaja analüüsida mitmeid erinevaid proove. Lisaks on suur võimalus, et ka optimeeritud tingimustel saab vale-positiivseid tulemusi, sest PCR-i mõjutavad väga mitmed tegurid.

Klassikaline Sanger sekveneerimine on samuti populaarne valik gRNA efektiivsuse kinnitamiseks. Praeguseks on arendatud mitmeid programme, mis võimaldavad analüüsida ka CRISPR/Cas9 transfekteeritud segapopulatsioonide proove. Kuna Sanger sekveneerimine on lihtne ja odav, on kirjanduses CRISPR/Cas9 tulemuste kinnitamiseks Sanger sekveneerimist kasutatud. Sanger sekveneerimisest saadud andmete analüüsil saab määravaks indelide osakaal järjestuses, andmete kvaliteet ja analüüsi programmides olevad piirangud (madalaim kvaliteediskoor, mis on programmis lubatud; failide formaat). Kui järjestused sisaldavad suure osakaaluga erinevaid indeleid, võib juhtuda, et need failid ei läbi programmides olevaid kvaliteedihinnanguid, mistõttu neid proove ei analüüsita. Ka antud töös tuli välja, et mõningatel juhtudel ei analüüsi Synthego ICE Sanger sekveneerimisest saadud faile. Synthego ICE andis mõningal juhul põhjuseks failide formaadi, väites, et analüüsitavad failid pole .ab1 formaadis või oli probleemiks liiga madal kvaliteedihinnang. Üldiselt on siiski Sanger sekveneerimisel põhinevad tulemused usaldusväärsemad kui PCR-i põhised meetodid, kuna Sanger sekveneerimise puhul saab hinnangu indelide positsioonile ja tüüpide kohta.

Kõige ulatuslikuma ja usaldusväärsema info rakupopulatsioonis esinevate indelide kohta annab siiski järgmise põlvkonna sekveneerimine. Selle meetodiga saab põhjaliku ülevaate järjestustes esinevate indelide positsiooni ja tüüpide kohta. Lisaks saab täpselt hinnata, kui palju oli indeleid sisaldavaid lugemeid ning kui paljud indelid viivad geeni funktsiooni rikkumiseni. PCR-i põhiste meetodite arenduse põhjenduseks on alati väited, et NGS on kallis, ajamahukas, andmete analüüs ei ole lihtne ning NGS ise ei ole kõigile kättesaadav. Tänapäevaks on enamus ülikoolidel ja teadusasutustele NGS-i kasutamise võimalus siiski olemas. NGS-i miinuseks on suhteliselt kõrge hind, seda eriti juhtudel, kui analüüsitakse üksikuid proove, ei ole võimalust proove multipleksida või teadusasutusel puudub NGS-i masin. Ka Tartu Ülikooli Genoomika Instituudi tuumiklaboritel on olemas kaks Illumina sekvenaatorit ning neilt on võimalus sekveneerimise teenust osta. Samuti

on võimalik sekveneerimise teenust osta erafirmadelt, kes tegelevad geneetiliste analüüsidega, seega ei tohiks NGS-i kättesaadavus ja hind olla suureks probleemiks. NGS on küll suhteliselt aeganõudev- olenevalt analüüsitavatest proovidest saab NGS-i puhul analüüsitavad andmed 3-5 päevaga. Samas arvestades, et PCR-i põhiseid meetodeid peab optimeerima iga proovi jaoks eraldi, muutuvad ka need meetodid ajakulult võrreldavaks. Eriti juhul kui tulemused ei ole korratavad ning kontroll proovid ei tööta. PCR-i puhul on tulemused tihti umbmääraseks ja põhinevad eeldusel, et kõik „töötavad nii nagu peab“. Seevastu NGS annab väga selged ja üheselt tõlgendatavad tulemused kõigi jaoks: kasutajal on võimalus näha indeli positsiooni, tüüpi ja sagedust. Selliselt saab valida testitud gRNA-dest kõige efektiivsema. NGS-i tulemustest ei jää ka kasutajale kahtlust, kas gRNA on töötanud või mitte. NGS-i andmete puhul on oluline andmete analüüs ning analüüsitavas programmis olevate parameetrite valik. Tüüpiliselt võib analüüsi käigus välja tulla, et ühe nukleotiidilisi muutusi leidub ka kontroll proovis. Nagu ka eelnevalt mainitud võivad sellised muutused toimuda juba PCR-i käigus, aga oma osa on ka sekveneerimisel tekkivatel vigadel. Sellisel juhul on oluline CRISPR/*Cas9* tulemuste korrektseks interpreteerimiseks võtta arvesse kontrollproovides esinevat vigade hulka.

Kokkuvõtvalt võib öelda, et võimalusel tuleks CRISPR/*Cas9* katse tulemuste hindamisel eelistada järgmise põlvkonna sekveneerimist või alternatiivselt Sanger sekveneerimist. Antud töös testitud PCR-i põhised meetodid ei anna gRNA efektiivsuse hindamisel alati selgelt interpreteeritavaid tulemusi, mille tõttu ei ole soovitatav neid kasutada.

KOKKUVÕTE

Praeguseks on gRNA aktiivsuse ja indel mutatsioonide sageduse hindamiseks välja töötatud mitmeid meetodeid, kuid kõigil neil on omad kitsaskohad, mistõttu ei anna need sageli usaldusväärseid tulemusi. Antud töös sai testitud kolme PCR-i põhist meetodit: ACT-PCR, Bindel-PCR ja ORNi PCR-i ning kahte sekveneerimise meetodit. Testitud PCR-i põhistest meetoditest töötas kõige efektiivsemalt ORNi PCR. Antud meetodiga saadud tulemused olid võrreldes teise kahe PCR-i põhiste meetoditega täpsemad ja selgemini tõlgendatavad. Kuna ACT- ja Bindel-PCR töötavad mõlemad alleel-spetsiifilisel PCR-i põhimõttel, alahinnatakse nende meetodite puhul indeleid sisaldavate järjestuste hulka kogu populatsioonis. Tulemuste analüüsimisel oli näha, et suure osa järjestusest moodustasid ka metsik-tüüpi ehk muundamata järjestused, mille puhul toimub edukas PCR. Seevastu ORNi PCR-i puhul blokeeritakse just metsik-tüüpi järjestused ning PCR-i amplifikatsioon toimub indeleid sisaldavatelt järjestustelt. Kõigi kolme meetodite puhul tõlgendatakse tulemusi vaid PCR bändi intensiivsuse põhjal, mis on umbmäärane ning ei anna selget informatsiooni gRNA efektiivsuse kohta. Võrreldes PCR-i põhiste meetoditega on Sanger sekveneerimine usaldusväärsem ja täpsem, kuna andmed sisaldavad infot nii indelite positsiooni kui ka tüübi kohta. Praeguseks on loodud mitmeid programme nagu Synthego ICE, mis võimaldab analüüsida Sanger sekveneerimisega ka transfekteeritud rakkude segapopulatsioone. Selliselt on võimalik saada oluliselt täpsemat informatsiooni indelite suuruse, sageduse ja positsiooni kohta. Antud töö tulemuste põhjal on näha, et gRNA efektiivsuse ja tekkinud indelite positsiooni kohta annavad kõige usaldusväärsemat infot Sanger- ja järgmise põlvkonna sekveneerimine. Samas mõjutavad ka Sanger sekveneerimise tulemuste analüüsi andmete kvaliteet ja indelite tüübid populatsioonis (indelite osakaalu või positsiooni tõttu võib Sanger sekveneerimisest saadud andmete kvaliteedihinnang olla madal). Programmides olevate piirangute tõttu ei ole võimalik Sanger sekveneerimisest saadud faile mõningatel juhtudel analüüsida. Kuna CRISPR/*Cas9* eksperimentides on vaja võimalikult täpset informatsiooni tulemuste kohta, on mõistlik kasutada sellisel juhul NGS-i. Ka antud töös selgus, et kõike täpsema ja ülevaatlikuma informatsiooni indelite sageduse kohta populatsioonis annab järgmise põlvkonna sekveneerimine.

Comparison of methods to evaluate gRNA efficiency using CRISPR/Cas9 technology for genome editing

Rita Kreevan

SUMMARY

Genome editing tools, such as TAL effector nucleases, zinc finger nucleases and RNA-guided endonucleases, enable targeted mutagenesis of selected DNA sequence in genomes of many species. The efficacy and mutation spectrum of genome editing methods can vary substantially depending on the target sequence. Once the CRISPR/Cas9 construct has been successfully delivered to the desired target, the efficiency and accuracy of the selected gRNA must be confirmed. Several methods have been developed to assess gRNA activity and the frequency of indel mutations, but they all have their limitations. Three PCR based methods were tested in this work: ACT-PCR, Bindel-PCR and ORNi PCR. Based on the results, ORNi PCR worked most efficiently amongst the three tested methods, giving results that are more accurate to interpret. Because ACT- and Bindel PCR are both based on allele-specific PCR, they underestimate the number of indel-containing sequences in whole cell populations. Analysis of the results showed that in all samples a large part of the sequences were wild type and these unmodified sequences will be amplified by the above mentioned methods. In contrast, in the case of ORNi PCR, the wild type sequences are blocked in the PCR reaction and only the sequences that contain indels are PCR amplified. In all three methods, the results are interpreted only on the basis of the PCR band intensity, which is vague and does not provide clear information about gRNA efficiency. Compared to the PCR-based methods, Sanger sequencing and NGS are more reliable and accurate. To date, several programs have been developed, such as Synthego ICE, which allow the analysis of mixed populations of transfected cells by Sanger sequencing. In this way, it is possible to obtain information about the size, frequency and position of the indels. Based on these results, it can be concluded that Sanger and next-generation sequencing give the most reliable information to assess the efficiency of gRNA and the position of the resulting indels. Due to the limitations that some programs have, it is not always possible to analyze files from Sanger sequencing. As NGS give you overall estimation of indels, it makes sense to use NGS to get the most reliable information about the outcome of the experiment. In addition, there is a wider range of different programs for analyzing data from NGS. The current work also confirmed that NGS provides the most accurate and comprehensive information about the indel frequencies in the cell population.

KIRJANDUSE LOETELU

- Allen, F., Crepaldi, L., Alsinet, C., Strong, A.J., Kleshchevnikov, V., De Angeli, P., Páleníková, P., Khodak, A., Kiselev, V., Kosicki, M., Bassett, A.R., Harding, H., Galanty, Y., Muñoz-Martínez, F., Metzakopian, E., Jackson, S.P., Parts, L., 2019. Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol* 37, 64–72.
- Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., Pul, Ü., 2014. Detection and characterization of spacer integration intermediates in type I-E CRISPR–Cas system. *Nucleic Acids Res* 42, 7884–7893.
- Ata, H., Ekstrom, T.L., Martínez-Gálvez, G., Mann, C.M., Dvornikov, A.V., Schaeffbauer, K.J., Ma, A.C., Dobbs, D., Clark, K.J., Ekker, S.C., 2018. Robust activation of microhomology-mediated end joining for precision gene editing applications. *PLoS Genet* 14, e1007652.
- Bell, C.C., Magor, G.W., Gillinder, K.R., Perkins, A.C., 2014. A high-throughput screening strategy for detecting CRISPR-Cas9 induced mutations using next-generation sequencing. *BMC Genomics* 15, 1002.
- Brinkman, E.K., Chen, T., Amendola, M., van Steensel, B., 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research* 42, e168–e168.
- Brinkman, E.K., van Steensel, B., 2019. Rapid Quantitative Evaluation of CRISPR Genome Editing by TIDE and TIDER, in: Luo, Y. (Ed.), *CRISPR Gene Editing*. Springer New York, New York, NY, pp. 29–44.
- Bruyneel, A.A.N., Colas, A.R., Karakikes, I., Mercola, M., 2019. AlleleProfileR: A versatile tool to identify and profile sequence variants in edited genomes. *PLoS ONE* 14, e0226694.
- Chakrabarti, A.M., Henser-Brownhill, T., Monserrat, J., Poetsch, A.R., Luscombe, N.M., Scaffidi, P., 2019. Target-Specific Precision of CRISPR-Mediated Genome Editing. *Molecular Cell* 73, 699–713.e6.
- Chandrasegaran, S., Carroll, D., 2016. Origins of Programmable Nucleases for Genome Engineering. *Journal of Molecular Biology* 428, 963–989.
- Choulika, A., Perrin, A., Dujon, B., Nicolas, J.F., 1995. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 15, 1968–1973.
- Deveau, H., Garneau, J.E., Moineau, S., 2010. CRISPR/Cas System and Its Role in Phage-Bacteria Interactions. *Annu. Rev. Microbiol.* 64, 475–493.

- Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E., 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* 32, 1262–1267.
- Donohoue, P.D., Barrangou, R., May, A.P., 2018. Advances in Industrial Biotechnology Using CRISPR-Cas Systems. *Trends in Biotechnology* 36, 134–146.
- Doudna, J.A., Charpentier, E., 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096.
- Farrar, J.S., Wittwer, C.T., 2017. High-Resolution Melting Curve Analysis for Molecular Diagnostics, in: *Molecular Diagnostics*. Elsevier, pp. 79–102.
- Findlay, S.D., Vincent, K.M., Berman, J.R., Postovit, L.-M., 2016. A Digital PCR-Based Method for Efficient and Highly Specific Screening of Genome Edited Cells. *PLoS ONE* 11, e0153901.
- Fujita, T., Yuno, M., Kitaura, F., Fujii, H., 2018. Detection of genome-edited cells by oligoribonucleotide interference-PCR. *DNA Research* 25, 395–407.
- Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* 109, E2579–E2586.
- Green, S.J., Venkatramanan, R., Naqib, A., 2015. Deconstructing the Polymerase Chain Reaction: Understanding and Correcting Bias Associated with Primer Degeneracies and Primer-Template Mismatches. *PLoS ONE* 10, e0128122.
- Gupta, D., Bhattacharjee, O., Mandal, D., Sen, M.K., Dey, D., Dasgupta, A., Kazi, T.A., Gupta, R., Sinharoy, S., Acharya, K., Chattopadhyay, D., Ravichandiran, V., Roy, S., Ghosh, D., 2019. CRISPR-Cas9 system: A new-fangled dawn in gene editing. *Life Sciences* 232, 116636.
- Gupta, R.M., Musunuru, K., 2014. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J. Clin. Invest.* 124, 4154–4161.
- Harayama, T., Riezman, H., 2017. Detection of genome-edited mutant clones by a simple competition-based PCR method. *PLoS ONE* 12, e0179165.
- Hensel, G., Kumlehn, J., 2019. Genome Engineering Using TALENs, in: Harwood, W.A. (Ed.), *Barley*. Springer New York, New York, NY, pp. 195–215.
- Hough, S.H., Ajetunmobi, A., Brody, L., Humphries-Kirilov, N., Perello, E., 2016. Desktop Genetics. *Personalized Medicine* 13, 517–521.

- Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., Liu, D.R., 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63.
- Hua, Y., Wang, C., Huang, J., Wang, K., 2017. A simple and efficient method for CRISPR/Cas9-induced mutant screening. *Journal of Genetics and Genomics* 44, 207–213.
- Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A.T., Charpentier, E., Nogales, E., Doudna, J.A., 2014. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* 343, 1247997–1247997.
- Kabadi, A.M., Ousterout, D.G., Hilton, I.B., Gersbach, C.A., 2014. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Research* 42, e147–e147.
- Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., Siksnys, V., 2013. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biology* 10, 841–851.
- Khanzadi, M.N., Khan, A.A., 2019. CRISPR/Cas9: Nature's gift to prokaryotes and an auspicious tool in genome editing. *J Basic Microbiol.*
- Kim, Hyojin, Um, E., Cho, S.-R., Jung, C., Kim, Hyongbum, Kim, J.-S., 2011. Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* 8, 941–943.
- Kosicki, M., Tomberg, K., Bradley, A., 2018. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* 36, 765–771.
- Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., Valen, E., 2019. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research* 47, W171–W174.
- Liu, C., Zhang, L., Liu, H., Cheng, K., 2017. Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications. *Journal of Controlled Release* 266, 17–26.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., van der Oost, J., Koonin, E.V., 2011. Evolution and classification of the CRISPR–Cas systems. *Nat Rev Microbiol* 9, 467–477.
- Mandal, P.K., Ferreira, L.M.R., Collins, R., Meissner, T.B., Boutwell, C.L., Friesen, M., Vrbanc, V., Garrison, B.S., Stortchevoi, A., Bryder, D., Musunuru, K., Brand, H., Tager, A.M., Allen, T.M., Talkowski, M.E., Rossi, D.J., Cowan, C.A., 2014. Efficient Ablation of Genes in Human Hematopoietic Stem and Effector Cells using CRISPR/Cas9. *Cell Stem Cell* 15, 643–652.

- Mock, U., Hauber, I., Fehse, B., 2016. Digital PCR to assess gene-editing frequencies (GEF-dPCR) mediated by designer nucleases. *Nat Protoc* 11, 598–615.
- Montgomery, J., Wittwer, C.T., Palais, R., Zhou, L., 2007. Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. *Nat Protoc* 2, 59–66.
- Park, J., Lim, K., Kim, J.-S., Bae, S., 2017. Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics* 33, 286–288.
- Pinello, L., Canver, M.C., Hoban, M.D., Orkin, S.H., Kohn, D.B., Bauer, D.E., Yuan, G.-C., 2016. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol*
- Qiu, P., Shandilya, H., D'Alessio, J.M., O'Connor, K., Durocher, J., Gerard, G.F., 2004. Mutation detection using SurveyorTM nuclease. *BioTechniques* 36, 702–707.
- Ramlee, M.K., Yan, T., Cheung, A.M.S., Chuah, C.T.H., Li, S., 2015. High-throughput genotyping of CRISPR/Cas9-mediated mutants using fluorescent PCR-capillary gel electrophoresis. *Sci Rep* 5, 15587.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281–2308.
- Robb, G.B., 2019. Genome Editing with CRISPR-Cas: An Overview. *Current Protocols Essential Laboratory Techniques* 19.
- Romay, G., Bragard, C., 2017. Antiviral Defenses in Plants through Genome Editing. *Front. Microbiol.* 8.
- Sakurai, T., Kamiyoshi, A., Takei, N., Watanabe, S., Sato, M., Shindo, T., 2019. Bindel-PCR: a novel and convenient method for identifying CRISPR/Cas9-induced biallelic mutants through modified PCR using *Thermus aquaticus* DNA polymerase. *Sci Rep* 9, 9923.
- Samarut, É., Lissouba, A., Drapeau, P., 2016. A simplified method for identifying early CRISPR-induced indels in zebrafish embryos using High Resolution Melting analysis. *BMC Genomics* 17, 547.
- Sfeir, A., Symington, L.S., 2015. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends in Biochemical Sciences* 40, 701–714.
- Shabbir, M.A.B., Hao, H., Shabbir, M.Z., Hussain, H.I., Iqbal, Z., Ahmed, S., Sattar, A., Iqbal, M., Li, J., Yuan, Z., 2016. Survival and Evolution of CRISPR–Cas System in Prokaryotes and Its Applications. *Front. Immunol.* 7.
- Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., Paques, F., 2011. Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy. *CGT* 11, 11–27.

- Sorek, R., Kunin, V., Hugenholtz, P., 2008. CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* 6, 181–186.
- Taheri-Ghahfarokhi, A., Taylor, B.J.M., Nitsch, R., Lundin, A., Cavallo, A.-L., Madeyski-Bengtson, K., Karlsson, F., Clausen, M., Hicks, R., Mayr, L.M., Bohlooly-Y, M., Maresca, M., 2018. Decoding non-random mutational signatures at Cas9 targeted sites. *Nucleic Acids Research* 46, 8417–8434.
- Ugozzoli, L., Wallace, R., 1991. Allele-specific polymerase chain reaction. *Methods* 2, 42–48.
- van Overbeek, M., Capurso, D., Carter, M.M., Thompson, M.S., Frias, E., Russ, C., Reece-Hoyes, J.S., Nye, C., Gradia, S., Vidal, B., Zheng, J., Hoffman, G.R., Fuller, C.K., May, A.P., 2016. DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Molecular Cell* 63, 633–646.
- Vouillot, L., Th  lie, A., Pollet, N., 2015. Comparison of T7E1 and Surveyor Mismatch Cleavage Assays to Detect Mutations Triggered by Engineered Nucleases. *G3* 5, 407–415.
- Wiedenheft, B., Sternberg, S.H., Doudna, J.A., 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338.
- Wilbie, D., Walther, J., Mastrobattista, E., 2019. Delivery Aspects of CRISPR/Cas for in Vivo Genome Editing. *Acc. Chem. Res.* 52, 1555–1564.
- Wiles, M.V., Qin, W., Cheng, A.W., Wang, H., 2015. CRISPR–Cas9-mediated genome editing and guide RNA design. *Mamm Genome* 26, 501–510.
- Wolfe, S.A., Neklodova, L., Pabo, C.O., 2000. DNA Recognition by Cys₂ His₂ Zinc Finger Proteins. *Annu. Rev. Biophys. Biomol. Struct.* 29, 183–212.
- Wood, A.J., Lo, T.-W., Zeitler, B., Pickle, C.S., Ralston, E.J., Lee, A.H., Amora, R., Miller, J.C., Leung, E., Meng, X., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Meyer, B.J., 2011. Targeted Genome Editing Across Species Using ZFNs and TALENs. *Science* 333, 307–307.
- Xu, H., Xiao, T., Li, W., Meyer, C.A., Wu, Q., Wu, D., Cong, L., Zhang, F., Liu, J.S., Brown, M., Liu, X.S., 2015. Sequence determinants of improved CRISPR sgRNA design. *Genome Res.* 25, 1147–1157.
- Yang, B., Wen, X., Kodali, N.S., Oleykowski, C.A., Miller, C.G., Kulinski, J., Besack, D., Yeung, J.A., Kowalski, D., Yeung, A.T., 2000. Purification, Cloning, and Characterization of the CEL I Nuclease[†]. *Biochemistry* 39, 3533–3541.
- Yu, C., Zhang, Y., Yao, S., Wei, Y., 2014. A PCR Based Protocol for Detecting Indel Mutations Induced by TALENs and CRISPR/Cas9 in Zebrafish. *PLoS ONE* 9, e98282.

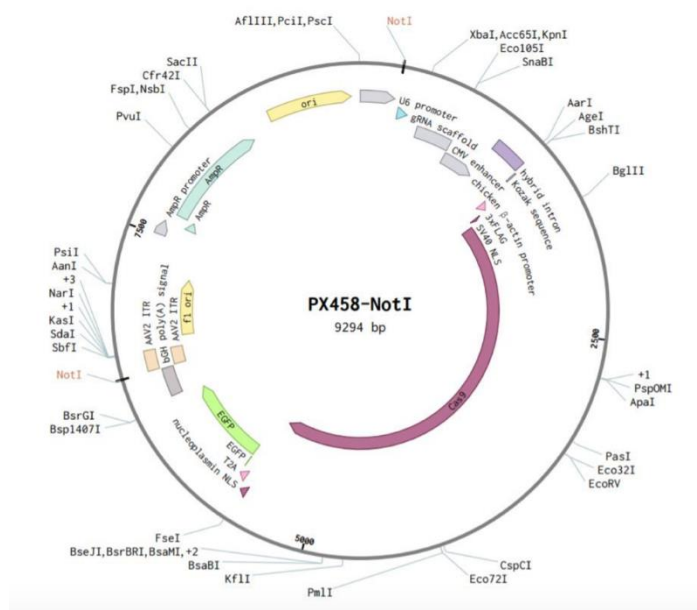
Zhu, X., Xu, Y., Yu, S., Lu, L., Ding, M., Cheng, J., Song, G., Gao, X., Yao, L., Fan, D., Meng, S., Zhang, X., Hu, S., Tian, Y., 2015. An Efficient Genotyping Method for Genome-modified Animals and Human Cells Generated with CRISPR/Cas9 System. *Sci Rep* 4, 6420.

LISAD

LISA 1. Erinevad meetodid CRISPR/Cas mutatsioonide kinnitamiseks

	Meetod	Tundlikkus	Ajakulu	Kommentaar	Viide
Heteroduplekside analüüsil põhinevad meetodid	T7E1	Madal	+	Sobib paremini pikemate indelite jaoks, kuid nukleotiidsse järjestuse kohta ei saa infot. Ei erista homosügootseid mutante.	(Mashal et al., 1995)
	Surveyor nuclease	Madal	+	Sobib paremini pikemate indelite jaoks, kuid nukleotiidsse järjestuse kohta ei saa infot. Ei erista homosügootseid mutante.	(Qiu et al., 2004)
	HRMA	Madal	++	Vajab spetsiifilist masinat, ei sobi pikemate indelite jaoks.	(Samarut et al., 2016)
	Heteroduplex mobility assay	Madal	+	Sobib paremini pikemate indelite jaoks, kuid nukleotiidsse järjestuse kohta ei saa infot. Ei erista homosügootseid mutante.	(Delwart et al., 1993)
Alleel-spetsiifilise PCR variatsioonid	cbPCR	Madal	+	Lihtne ja odav, aga ei ole kvantitatiivne.	(Harayama and Riezman, 2017)
	Bindel-PCR	Madal	++	Lihtne ja kiire.	(Sakurai et al., 2019)
	ACT-PCR	Madal	+	Lihtne ja kiire. Vajab ainult mutatsiooni spetsiifilist praimerit ja termotsüklerit.	(Hua et al., 2017)
	ORNi-PCR	Madal	+	Vaja RNA oligoribonukleotiide.	(Fujita et al., 2018)
	ssdPCR	Madal	+	Vajab spetsiifilist masinat.	(Falabella et al., 2017)
	MSBSP-PCR	Madal	+	Mutatsioonide tuvastamine on piiratud.	(Guo et al., 2018)
Fluorestseeruvate sondidega PCR-id	ddPCR	Hea	+	Vajab spetsiifilist masinat.	(Findlay et al., 2016)
	GEF-dPCR	Hea	++	Vajab spetsiifilist masinat.	(Mock et al., 2016)
	PCR-LDR	Madal	+	Vaja mitu paari sonde erinevate mutatsioonide jaoks.	(Kc et al., 2016)
	qEva-CRISPR	Madal	+	Mutatsioonide omaduste kohta ei saa infot.	(Dabrowska et al., 2018)

Sekveneerimisel põhinevad meetodid	IDAA	Hea	++	Vajab PCR-i masinat	(Yang et al., 2015)
	CRES-seq	Hea	+++	Analüüsiks vaja Illumina masinat.	(Veeranagouda et al., 2018)
	Sanger sekveneerimine	Väga hea	+++	Töömahukas, aga saab infot indelite suuruse ja järjestuse kohta.	(Brinkman and van Steensel, 2019; Yang et al., 2015)
	NGS	Väga hea	+++	Töömahukas, aga saab infot indelite suuruse ja järjestuse kohta.	(Bell et al., 2014)



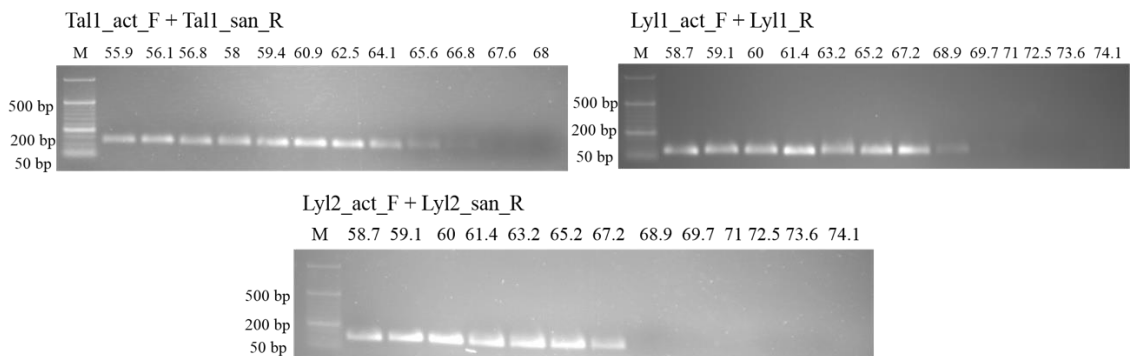
LISA 2. pX458-NotI plasmiid

LISA 3. NGS, Sanger sekveneerimise ja ORNi PCR-is kasutatud praimerid

Lyl1, Lyl2	<i>Lyl_san_F</i>	5' – GCCCTCCTTTTAGGTCCTAGC – 3'
	<i>Lyl_san_R</i>	5' – CAGGGAGTGCACGTGTTTC – 3'
Tal1	<i>Tal_san_F</i>	5' – GGAGGGACTGGGTGATCTTTC – 3'
	<i>Tal_san_R</i>	5' – CGTTGAGCAGGACTAGGTGC – 3'

LISA 4. ACT-PCR-i jaoks kasutatud praimerid

Tal1	<i>Tal_act_F</i>	5' – ACCCTGTTACAGGATGACGGA – 3'
	<i>Tal_san_R</i>	5' – CGTTGAGCAGGACTAGGTGC – 3'
Lyl1	<i>Ly_actl_F</i>	5' – GGGCGGGCACATGGACCCCA – 3'
	<i>Lyl_R</i>	5' – AAAGCTCAGGTAAACAGGTGG – 3'
Lyl2	<i>Lyl2_act_F</i>	5' – CGTGGGGTCCATGTGCCCCGCC – 3'
	<i>Lyl2_san_R</i>	5' – CAGGGAGTGCACGTGTTTC – 3'

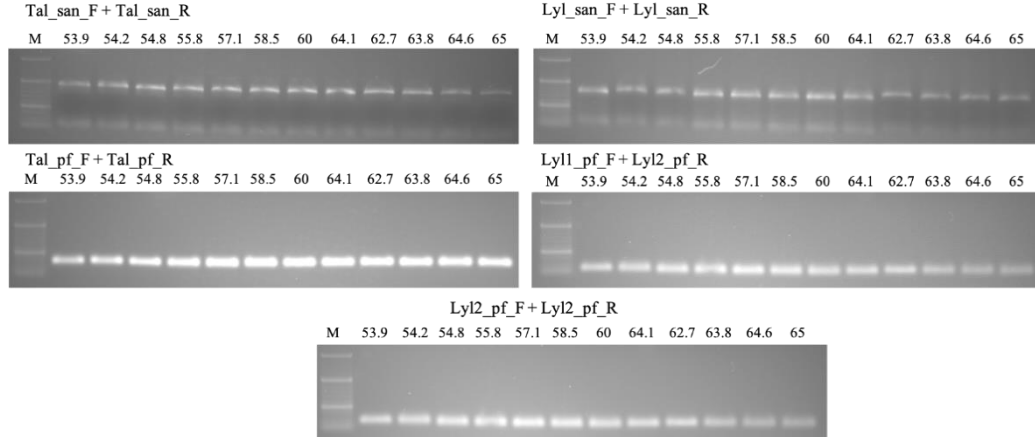


LISA 5. ACT - PCR-i optimiseerimine. Praimerite sobiva sulamistemperatuuri testimiseks viidi läbi gradient PCR CRISPR-Tal1 jaoks vahemikus 55,9°C - 68°C, CRISPR-Lyl1 jaoks vahemikus 58,7°C - 74,1°C ja CRISPR-Lyl2 jaoks vahemikus 58,7°C - 74,1°C Vastavalt sellele valiti välja igale praimerile sobiv sulamistemperatuur. CRISPR-Tal1 praimerite sobiv temperatuur oli 66,8°C CRISPR-Lyl1 ja CRISPR-Lyl2 puhul 67,2°C

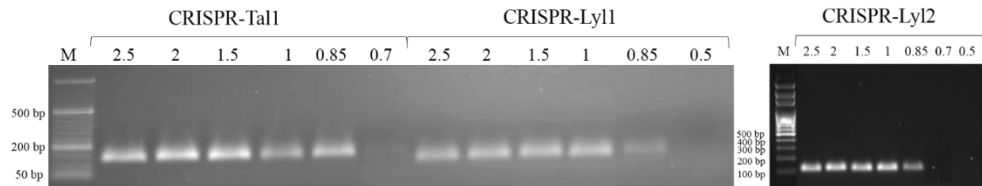
LISA 6. Bindel-PCRis kasutatud praimerid

Tal1	<i>Tal_san_F</i>	5' – GGAGGGACTGGGTGATCTTTC – 3'
	<i>Tal_san_R</i>	5' – CGTTGAGCAGGACTAGGTGC – 3'
	<i>Tal_pf_F</i>	5' – ACCCTGTTACAGGATGACGG – 3'
	<i>Tal_pf_R</i>	5' – GCCGGCTCCGGGGCATGGC – 3'
Lyl1	<i>Lyl_san_F</i>	5' – GCCCTCCTTTTAGGTCCTAGC – 3'
	<i>Lyl_san_R</i>	5' – CAGGGAGTGCACGTGTTTC – 3'
	<i>Lyl_pf_F</i>	5' – GGGCGGGCACATGGACCCCA – 3'
	<i>Lyl_pf_R</i>	5' – AAAGCTCAGGTAAACAGGTGG – 3'
Lyl2	<i>Lyl_san_F</i>	5' – GCCCTCCTTTTAGGTCCTAGC – 3'
	<i>Lyl_san_R</i>	5' – CAGGGAGTGCACGTGTTTC – 3'
	<i>Lyl2_pf_F</i>	5' – TGGGGTCCATGTGCCCCGCC – 3'
	<i>Lyl2_pf_R</i>	5' – CTCTGTAGAAAGGGGCCAG – 3'

A



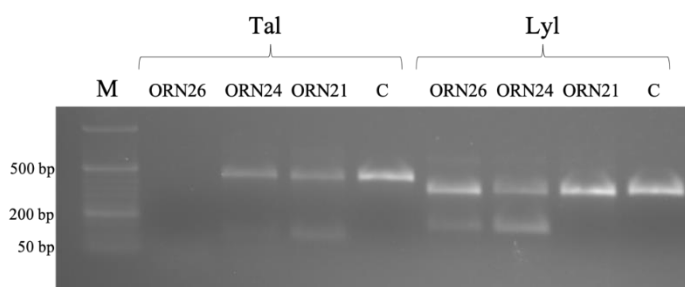
B



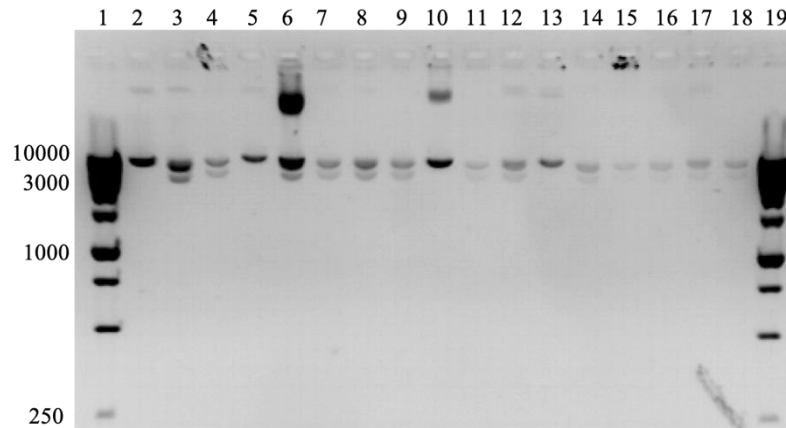
LISA 7. Bindel-PCR praimerite ja MgCl₂ optimeerimine. **A** Praimerite sobiva sulamistemperatuuri testimiseks viidi läbi gradient PCR temperatuuride vahemikus 53,9°C -65°C. Vastavalt sellele valiti igale praimerile paarile sobiv temperatuur. Tal1 san praimerite sulamistemperatuuriks jäi 60°C ja Lyl1 ja Lyl2 puhul 57,1°C . Tal1, Lyl1 ja Lyl2 pf praimerite temperatuuriks jäi 57°C. **B** Sobiva MgCl₂ kontsentratsiooni määramine pf praimeritel. Madalaima MgCl₂ määramiseks testiti MgCl₂ kontsentratsioone vahemikus 2.5 – 0.7 mM. Kõikide testitud praimerite puhul määrati MgCl₂ kontsentratsiooniks 0.85 mM.

LISA 8. ORNi PCR-i jaoks disainitud oligonukleotiidid

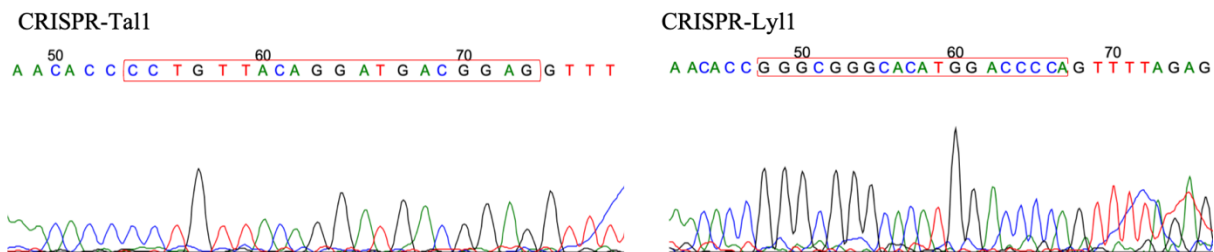
Tal1	ORN 26	5' – CTGTTACAGGATGACGGAGCGGCCG – 3'
	ORN 24	5' – CTGTTACAGGATGACGGAGCGGCC – 3'
	ORN 20	5' – CAGGATGACGGAGCGGCCG – 3'
Lyl1, Lyl2	ORN 26	5' – ATGUGCCCGCCCCAGGCCAGAGCAGA – 3'
	ORN 24	5' – GCCCGCCCCAGGCCAGAGCAGAGG – 3'
	ORN 20	5' – ATGTGCCCGCCCCAGGCCAG – 3'



LISA 9. ORNi PCR-i oligonukleotiidide testimine. Iga CRISPR konstrukti jaoks testiti kolme erineva pikkusega oligonukleotiidi: 26 nukleotiidiline, 24 ja 21 nukleotiidilised. CRISPR-Tal1 puhul töötas kõige efektiivsemalt ORN26, mis blokeeris täielikult PCR-i. CRISPR-Lyl1 puhul ei blokeerinud ükski testitud oligonukleotiididest PCR-i täielikult, kuid osaliselt töötas kõige paremini ORN24.



LISA 10. CRISPR-Lyl1 ja CRISPR-Lyl2 plasmiidide lõikamine *NotI* restriksiooniensüümiga. 1 rada- Fermentas O'GeneRuler 1 kb DNA marker; rada 2-10 - restrikteeritud CRISPR-Lyl1 kolooniad; 11-18 - restrikteeritud CRISPR-Lyl2 kolooniad 19 - Fermentas O'GeneRuler 1 kb DNA marker. Kui plasmidi on inserteerunud disainitud oligonukleotiid kaob järjestusest teine *NotI* lõikejärjestus ning *NotI* restriksiooniensüümiga lõigates plasmid lineariseerub ning tekib ~9 kb fragment. Kui insert ei ole korrektselt inserteerunud, lõigatakse plasmidi kahest kohast ja tekivad ~3 kb ja ~6 kb pikkused fragmendid. Õige konstruktid on proovides nr 1, 5, 10 ja 13.



LISA 11. CRISPR-Tal1 ja CRISPR-Lyl1 plasmiidide kontroll Sanger sekveneerimisega. Punasega on märgitud gRNA järjestus.

LISA 12. Ülevaade analüüsitavatest proovides ning analüüsi meetoditest

Rakud	Proov	Illumina NGS andmeanalüüs	Sanger sekveneerimise andmeanalüüs
EGFP+ segapopulatsioon	CRISPR-Tal1 JP	CrispRVariant	Synthego ICE
	CRISPR-Lyl1 JP		
	CRISPR-Lyl2 JP		
EGFP- segapopulatsioon	CRISPR-Tal1 JP-	CrispRVariant	Synthego ICE
	CRISPR-Lyl1 JP-		
	CRISPR-Lyl2 JP-		
EGFP+ segapopulatsioon	CRISPR-Tal1 JO	CrispRVariant	Synthego ICE
	CRISPR-Lyl1 JO		
	CRISPR-Lyl2 JO		
CRISPR-Tal1 EGFP+ kloonid	Tal1-1	CrispRVariant	Synthego ICE
	Tal1-2		
	Tal1-3		
	Tal1-4		
CRISPR-Lyl2 EGFP+ kloonid	Lyl2-1	CrispRVariant	Synthego ICE
	Lyl2-2		
	Lyl2-3		
	Lyl2-4		

LISA 13. FACS-i sorteerimise tulemused

Kasutatud reagent	Proovid	EGFP+ rakkude arv	EGFP- rakkude arv	Kokku	EGFP+%	EGFP-%
jetPRIME®	CRISPR-Tal1	17357	78709	96066	18%	81%
	CRISPR-Lyl1	28925	47714	76639	38%	62%
	CRISPR-Lyl2	40708	37378	78085	52%	48%
jetPRIME®	CRISPR-Tal1	7907	122699	130606	6%	94%
	CRISPR-Lyl1	21319	136067	153786	13%	86%

	CRISPR-Lyl2	90911	105399	196310	46%	54%
jetPRIME®	CRISPR-Tal1	4942	388286	393228	1%	99%
	CRISPR-Lyl1	24243	94579	118822	20%	80%
	CRISPR-Lyl2	25719	342964	368683	7%	93%
jetPRIME®	CRISPR-Tal1	96062	73544	82606	11%	89%
	CRISPR-Lyl1	61904	223251	285155	22%	78%
	CRISPR-Lyl2	41144	227388	268532	15%	85%
jetOPTIMUS®	CRISPR-Tal1	5060	116831	121891	4%	96%
	CRISPR-Lyl1	7606	79627	87233	9%	91%
	CRISPR-Lyl2	19344	126247	155591	13%	87%

LIHTLITSENTS

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Rita Kreevan

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

„Meetodite võrdlus gRNA efektiivsuse hindamiseks genoomi editeerimisel CRISPR/*Cas9* tehnoloogiat kasutades“, mille juhendaja on Tõnis Org PhD

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Rita Kreevan

11.06.2020